



BOARD OF REGENTS THE UNIVERSITY OF TEXAS SYSTEM (US)

FEATURES		source		Location/Qualifiers		1..3826		/organism="Homo sapiens"		/mol_type="unassigned DNA"		/db_xref="taxon:9606"	
ORIGIN		Query Match		100.0%; Score 3826; DB 6; Length 3826;		Best Local Similarity 100.0%; Pred. No. 0;		Matches 3826; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
QY	1	AGGCTGCGCT	GACGACGAGCGCT	GGCTGCT	TAAGCTCGCGGGGT	TAAGGGT	CGCGCTCGCGCTGGG	60					
DB	1	AGGCTGCGCT	GACGACGAGCGCT	GGCTGCT	TAAGCTCGCGGGGT	TAAGGGT	CGCGCTCGCGCTGGG	60					
QY	61	CCAGGGTTT	TGGGGCGCGGAT	CCGGCAGCT	GTAGCGGGGCGG	CACCCCTCTCTCT	TCTCTGCTGC	120					
DB	61	CCAGGGTTT	TGGGGCGCGGAT	CCGGCAGCT	GTAGCGGGGCGG	CACCCCTCTCTCT	TCTCTGCTGC	120					
QY	121	GGTCACAGC	CAATGTAC	CGGCTCGGCT	TGGCTGCCCCCT	TCCCCAGGAT	TCCCCATCCCA	180					
DB	121	GGTCACAGC	CAATGTAC	CGGCTCGGCT	TGGCTGCCCCCT	TCCCCAGGAT	TCCCCATCCCA	180					
QY	181	GCTTCTCGCC	TCCCGCAGCG	CCCCACCCCGG	ATTTGACACCCCT	TAAAGGCTCCAC	240						
DB	181	GCTTCTCGCC	TCCCGCAGCG	CCCCACCCCGG	ATTTGACACCCCT	TAAAGGCTCCAC	240						
QY	241	CCCGCTCGGG	ATCCCTTCTCC	AGCTCTAT	CCCTTAGGACT	GGCCCGCCCTAGAA	300						
DB	241	CCCGCTCGGG	ATCCCTTCTCC	AGCTCTAT	CCCTTAGGACT	GGCCCGCCCTAGAA	300						
QY	301	CCTCCCGT	CAGGATCTCG	TGCTCCTC	AGCGCTC	ACAGGCTCTCC	CGAGCGCCATCGCC	360					
DB	301	CCTCCCGT	CAGGATCTCG	TGCTCCTC	AGCGCTC	ACAGGCTCTCC	CGAGCGCCATCGCC	360					
QY	361	TTGAGCTGCC	CACCTACCTCTAG	ACTGCTCC	CGGGCTGGCGT	CCACGAGTCTCAGCC	420						
DB	361	TTGAGCTGCC	CACCTACCTCTAG	ACTGCTCC	CGGGCTGGCGT	CCACGAGTCTCAGCC	420						
QY	421	GGCGACCC	CTTCTCGCGTT	ACCTCTTC	CGGACAGCAC	CCCCCTCTCTCT	CTCGGTAGC	480					
DB	421	GGCGACCC	CTTCTCGCGTT	ACCTCTTC	CGGACAGCAC	CCCCCTCTCTCT	CTCGGTAGC	480					
QY	481	TCCTACCC	CTGCTGTGGGGCT	CGTCCCGG	CGCGCAGCCCT	CGGTGTGCTCTCCG	CAC	540					
DB	481	TCCTACCC	CTGCTGTGGGGCT	CGTCCCGG	CGCGCAGCCCT	CGGTGTGCTCTCCG	CAC	540					
QY	541	GGCGCGCG	CTCTCAGCG	CCCCCTG	CGCCCTCGGG	CCCCCTCTCTCT	CTGCTCGCCCTGAC	600					
DB	541	GGCGCGCG	CTCTCAGCG	CCCCCTG	CGCCCTCGGG	CCCCCTCTCTCT	CTGCTCGCCCTGAC	600					
QY	601	GCCATGGC	GTGAGCCTCA	AGGACGAGT	GTGTGCTCCAT	CTGCTGAGCAT	TACACG	660					
DB	601	GCCATGGC	GTGAGCCTCA	AGGACGAGT	GTGTGCTCCAT	CTGCTGAGCAT	TACACG	660					
QY	661	GACCGGTG	AGCCTGGGCTG	CGAGCACT	TTCTGCGCGCGT	TGATCAGGAGCACTGG	720						
DB	661	GACCGGTG	AGCCTGGGCTG	CGAGCACT	TTCTGCGCGCGT	TGATCAGGAGCACTGG	720						
QY	721	GTGCGC	AGGAGCG	CGCGCCCG	CGACTGCCCC	AGTGTGCGGCGCAC	GTTCGCCGAG	780					
DB	721	GTGCGC	AGGAGCG	CGCGCCCG	CGACTGCCCC	AGTGTGCGGCGCAC	GTTCGCCGAG	780					
QY	781	CCCGCGT	TGGCGCCAG	CTC	TAAGTGG	CAACATG	TGTGGAGCGCTAC	840					
DB	781	CCCGCGT	TGGCGCCAG	CTC	TAAGTGG	CAACATG	TGTGGAGCGCTAC	840					
QY	841	CTGAGCGC	ATCTCA	ACGCGCG	CGCGCGCG	CGACCTG	CCGACGCGCA	900					
DB	841	CTGAGCGC	ATCTCA	ACGCGCG	CGCGCGCG	CGACCTG	CCGACGCGCA	900					
QY	901	AAGCTT	CTTCTG	CTACG	ACCGCG	CGCTTCTCT	CTGCTTCTG	960					

DB	901	AAGCTT	CTTGC	CTAC	GGAC	CCGG	CGCT	CTCT	CTG	960	
QY	961	CACGAG	CAGCAT	CAGT	CACCGG	CATCG	AGCAG	CGCT	TCG	1020	
DB	961	CACGAG	CAGCAT	CAGT	CACCGG	CATCG	AGCAG	CGCT	TCG	1020	
QY	1021	AAGGAC	CAACTT	CAGG	CCCTT	CAAG	CAGC	AGCG	GGAACA	1080	
DB	1021	AAGGAC	CAACTT	CAGG	CCCTT	CAAG	CAGC	AGCG	GGAACA	1080	
QY	1081	CTTAAG	CGACAA	CTCGG	AGACCA	AGTCTT	CCAA	AGAG	CCTCGG	1140	
DB	1081	CTTAAG	CGACAA	CTCGG	AGACCA	AGTCTT	CCAA	AGAG	CCTCGG	1140	
QY	1141	GAGG	CCCTT	CGAG	CGGCT	GCAC	CGGCT	GCAG	CGGCA	1200	
DB	1141	GAGG	CCCTT	CGAG	CGGCT	GCAC	CGGCT	GCAG	CGGCA	1200	
QY	1201	CTGAGG	CGGGA	CAAC	CGGCG	CGAC	CGT	GAC	CGAT	1260	
DB	1201	CTGAGG	CGGGA	CAAC	CGGCG	CGAC	CGT	GAC	CGAT	1260	
QY	1261	CAGCAG	CTGCG	CAAG	GGTCC	AGGAG	GGCG	CGA	GAT	1320	
DB	1261	CAGCAG	CTGCG	CAAG	GGTCC	AGGAG	GGCG	CGA	GAT	1320	
QY	1321	GACCGG	CAACCT	CTCG	TGGG	GGTGG	GGCT	CACT	GT	1380	
DB	1321	GACCGG	CAACCT	CTCG	TGGG	GGTGG	GGCT	CACT	GT	1380	
QY	1381	CATGAG	CAACCA	CTCAT	TATGA	AGACTT	CCCG	AGCT	CCAA	1440	
DB	1381	CATGAG	CAACCA	CTCAT	TATGA	AGACTT	CCCG	AGCT	CCAA	1440	
QY	1441	TAACCA	ATCTG	GAAG	TCCCT	GTTC	AGAG	CAT	CCAG	1500	
DB	1441	TAACCA	ATCTG	GAAG	TCCCT	GTTC	AGAG	CAT	CCAG	1500	
QY	1501	GACCGG	CAACCG	CCAG	CGCT	GTAT	CTGT	CGAG	CGCT	1560	
DB	1501	GACCGG	CAACCG	CCAG	CGCT	GTAT	CTGT	CGAG	CGCT	1560	
QY	1561	GGCAACT	TGAC	CCAC	CAG	CCACT	GAG	CACT	CCCA	1620	
DB	1561	GGCAACT	TGAC	CCAC	CAG	CCACT	GAG	CACT	CCCA	1620	
QY	1621	GTGCTG	GGTCT	TAAG	CCCT	TAAG	TCCT	AGT	GGCT	1680	
DB	1621	GTGCTG	GGTCT	TAAG	CCCT	TAAG	TCCT	AGT	GGCT	1680	
QY	1681	AAGAC	CCAGT	GTGAT	CGGG	CTGG	CAAC	AGCG	CGCA	1740	
DB	1681	AAGAC	CCAGT	GTGAT	CGGG	CTGG	CAAC	AGCG	CGCA	1740	
QY	1741	ATCCAG	CCCG	CAG	CGCG	CTTCT	ACT	GCAT	CGT	1800	
DB	1741	ATCCAG	CCCG	CAG	CGCG	CTTCT	ACT	GCAT	CGT	1800	
QY	1801	TGCA	CGAG	CCCT	TGAC	CGG	CGCT	TAAC	CGT	1860	
DB	1801	TGCA	CGAG	CCCT	TGAC	CGG	CGCT	TAAC	CGT	1860	
QY	1861	CTGG	ACTAT	GAC	CAAG	CGT	TGCT	CTAT	CTT	1920	
DB	1861	CTGG	ACTAT	GAC	CAAG	CGT	TGCT	CTAT	CTT	1920	
QY	1921	ACCT	TCG	CGA	GAAG	TTCC	CTG	CTCT	TTA	1980	
DB	1921	ACCT	TCG	CGA	GAAG	TTCC	CTG	CTCT	TTA	1980	
QY	1981	GGCA	ATG	CAAG	CAAG	CTT	CAG	CGCT	CGG	2040	

Db 1981 GCCAATGGCAAGACGTTTCAGCGCTGCGGATCAACACCGTCCGCATCTAGTCCAGGCAG 2040  
Qy 2041 AAGGAGACCAACCTCTGGGACCACTGCACCTGCAGAGCCCTGCCAGGAGATAG 2100  
Db 2041 AAGGAGACCAACCTCTGGGACCACTGCACCTGCAGAGCCCTGCCAGGAGATAG 2100  
Qy 2101 AAGACCTGGAATCCAGGCCACCGTGGCCACTGGAGACCTCAGGCGAGTTGTTTACCTCC 2160  
Db 2101 AAGACCTGGAATCCAGGCCACCGTGGCCACTGGAGACCTCAGGCGAGTTGTTTACCTCC 2160  
Qy 2161 AGCCTCAGTCTGTAATAAGAGGTTGCAATCCCTTACTTCCCTTAACTCTCTCCAGCATC 2220  
Db 2161 AGCCTCAGTCTGTAATAAGAGGTTGCAATCCCTTACTTCCCTTAACTCTCTCCAGCATC 2220  
Qy 2221 GATGTTCTGTAGCTCTGACCTTGATAGGATACAGCTTTGATCCCAAGATGTGACATGGC 2280  
Db 2221 GATGTTCTGTAGCTCTGACCTTGATAGGATACAGCTTTGATCCCAAGATGTGACATGGC 2280  
Qy 2281 TTCTCCTCAGGGCAACCCCTGCCAACCCCTCATCCCATCTTCTCAGGGGAGGGGACTA 2340  
Db 2281 TTCTCCTCAGGGCAACCCCTGCCAACCCCTCATCCCATCTTCTCAGGGGAGGGGACTA 2340  
Qy 2341 CCTTCCAGTGTCTCCTCCAGCCAGCCCTGACCTCAGGAAAGTGTACAGATGCCAGT 2400  
Db 2341 CCTTCCAGTGTCTCCTCCAGCCAGCCCTGACCTCAGGAAAGTGTACAGATGCCAGT 2400  
Qy 2401 AGTTGGGAGCCGGAAGACACACAGCACCTCTTATGTCCCATGGCTTAAGACTTACCCC 2460  
Db 2401 AGTTGGGAGCCGGAAGACACACAGCACCTCTTATGTCCCATGGCTTAAGACTTACCCC 2460  
Qy 2461 TGACCAAGCTAGTGTGATGGCCATTTACCCCTTGACCCAGTCCACAGTGGTCAACAGT 2520  
Db 2461 TGACCAAGCTAGTGTGATGGCCATTTACCCCTTGACCCAGTCCACAGTGGTCAACAGT 2520  
Qy 2521 ACCTGGTCTTAGGGTTGCTGAGAGCCAACTCTCTGTCGCCACCCCAACCAAGAACTAT 2580  
Db 2521 ACCTGGTCTTAGGGTTGCTGAGAGCCAACTCTCTGTCGCCACCCCAACCAAGAACTAT 2580  
Qy 2581 ATGGTTCCTACTCTCCACCTGATCTGCTGCTGATGATGATGCTGCTGCTGCTGCTGCTG 2640  
Db 2581 ATGGTTCCTACTCTCCACCTGATCTGCTGCTGATGATGATGCTGCTGCTGCTGCTGCTG 2640  
Qy 2641 CACCTGGTGTAGTGCACACATATAGTATGTCGCCACCCACTTCTGCTGCCACAGGCC 2700  
Db 2641 CACCTGGTGTAGTGCACACATATAGTATGTCGCCACCCACTTCTGCTGCCACAGGCC 2700  
Qy 2701 GAGGACAGGTTAGGGTATACCAAGCTGATGATGATGATGATGATGATGATGATGATGATG 2760  
Db 2701 GAGGACAGGTTAGGGTATACCAAGCTGATGATGATGATGATGATGATGATGATGATGATG 2760  
Qy 2761 GCAGGACAGGCTCCCTGGATGATCGAGGTCCTGAGTCTGCAACAGAGTCCAGCCA 2820  
Db 2761 GCAGGACAGGCTCCCTGGATGATCGAGGTCCTGAGTCTGCAACAGAGTCCAGCCA 2820  
Qy 2821 ACCCTCTTACGCCAGGCTCTGTGACCTGTAGGTTGAGGAGGCTTCCAGAGCAGTTG 2880  
Db 2821 ACCCTCTTACGCCAGGCTCTGTGACCTGTAGGTTGAGGAGGCTTCCAGAGCAGTTG 2880  
Qy 2881 TTGTAATTAGACCAAGCACTGGAGGGCTGTGCTGCTAGACCCCTTGTGCTGCTGCTGCTG 2940  
Db 2881 TTGTAATTAGACCAAGCACTGGAGGGCTGTGCTGCTAGACCCCTTGTGCTGCTGCTGCTG 2940  
Qy 2941 ATCTATCTCAGTTAGGATCTGCTGCAAGAAACAGAGCCACTGTAGCTGTTTAAATTA 3000  
Db 2941 ATCTATCTCAGTTAGGATCTGCTGCAAGAAACAGAGCCACTGTAGCTGTTTAAATTA 3000  
Qy 3001 GACAAGGATTTACTACCTGGCCCTTGGTGGCTTGCMAAATTTGTTGGAAGAGCTGAGAAG 3060  
Db 3001 GACAAGGATTTACTACCTGGCCCTTGGTGGCTTGCMAAATTTGTTGGAAGAGCTGAGAAG 3060  
Qy 3061 CAGACTCTGCTGAATTTCCAGGAACTCCACGCGCAGATTCATCATGTCTGTGTGACCA 3120  
Db 3061 CAGACTCTGCTGAATTTCCAGGAACTCCACGCGCAGATTCATCATGTCTGTGTGACCA 3120

Qy 3121 GGAAAGCTGCCCATCTGCGAGGAGCCACTATGCCAGAAAGCTGCTGACTGCAGAACTA 3180  
Db 3121 GGAAAGCTGCCCATCTGCGAGGAGCCACTATGCCAGAAAGCTGCTGACTGCAGAACTA 3180  
Qy 3181 GGCTCCCTCTGCCACGGTCCGTGCCAGCAATAGATGTCCTGAGGCTGCCCTCTCCCA 3240  
Db 3181 GGCTCCCTCTGCCACGGTCCGTGCCAGCAATAGATGTCCTGAGGCTGCCCTCTCCCA 3240  
Qy 3241 CTTCACTCAGTTCCTCAAAATCTAAATTTTTTA CAAGAGATCTGTTTGGGGAACTTAAAGTC 3300  
Db 3241 CTTCACTCAGTTCCTCAAAATCTAAATTTTTTA CAAGAGATCTGTTTGGGGAACTTAAAGTC 3300  
Qy 3301 AGATCCAGAACTCTGGCTGCAAGGAGTCTGGGAAATGCTATTTCCTAGAGGAGTTA 3360  
Db 3301 AGATCCAGAACTCTGGCTGCAAGGAGTCTGGGAAATGCTATTTCCTAGAGGAGTTA 3360  
Qy 3361 GGGTGGGTGGAGCAAGCCCACTGCCGTTTTTCTGCCACAGCATCCAACTCGTGAAGAACT 3420  
Db 3361 GGGTGGGTGGAGCAAGCCCACTGCCGTTTTTCTGCCACAGCATCCAACTCGTGAAGAACT 3420  
Qy 3421 CGGAGAGGGTGGAGTCCACATCTAGGGTTGTCCTGCCCTTGGCTCTATCCCTGCCCAG 3480  
Db 3421 CGGAGAGGGTGGAGTCCACATCTAGGGTTGTCCTGCCCTTGGCTCTATCCCTGCCCAG 3480  
Qy 3481 AGGTGGGAACTGGAGGAGTGGGCTGCAAGACTGAGCTTAAATGCTCCCGGCTTGAATC 3540  
Db 3481 AGGTGGGAACTGGAGGAGTGGGCTGCAAGACTGAGCTTAAATGCTCCCGGCTTGAATC 3540  
Qy 3541 TTTCTTTCTAGTCTCGGGGCTTAGATCTGCACATTTGGGGTCTCTGACACACACACCATC 3600  
Db 3541 TTTCTTTCTAGTCTCGGGGCTTAGATCTGCACATTTGGGGTCTCTGACACACACACCATC 3600  
Qy 3601 CCAAGTAGCCGGAAGAGCTAAACACAGGGGGTCTTAAATGGTGTCCCGGCGCACCCG 3660  
Db 3601 CCAAGTAGCCGGAAGAGCTAAACACAGGGGGTCTTAAATGGTGTCCCGGCGCACCCG 3660  
Qy 3661 GGCTCCCTGGGCAAGGAAATGTCAGCCCTACCCCAACCCCTCACTACCACTACCAATCT 3720  
Db 3661 GGCTCCCTGGGCAAGGAAATGTCAGCCCTACCCCAACCCCTCACTACCACTACCAATCT 3720  
Qy 3721 GGGCCACCCAGCAGTATTTTATTTAAATGTTGGCCATTTTATGATGATGATCAAT 3780  
Db 3721 GGGCCACCCAGCAGTATTTTATTTAAATGTTGGCCATTTTATGATGATGATCAAT 3780  
Qy 3781 TGTATTAAATTAAGTTACAGATGTCATAAAAAAATAAAAAAATAAAAAA 3826  
Db 3781 TGTATTAAATTAAGTTACAGATGTCATAAAAAAATAAAAAAATAAAAAA 3826

RESULT 2  
AX877131 LOCUS AX877131 3436 bp DNA linear PAT 17-DEC-2003  
DEFINITION Sequence 12036 from Patent EP1074617.  
ACCESSION AX877131  
VERSION AX877131.1 GI:40031867  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
AUTHORS Ota, T., Isozaki, T., Nishikawa, T., Hayashi, K., Saiko, K., Yamamoto, J.,  
Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Otsuki, T.  
TITLE Primers for synthesising full-length cDNA and their use  
JOURNAL Patent: EP 1074617-A 12036 07-FEB-2001;  
Research Association for Biotechnology (JP)  
FEATURES  
source Location/Qualifiers  
1. 3436  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
235..1662 CDS

```

/note="unnamed protein product"
/codon_start=1
/protein_id="CAE89585.1"
/db_xref="GI:40031868"
/translation="MACSLKDELCSICLSIYQDPVSLGCBHYFRRRCITEHWVQEA
QCARDPCRTFAFPALAPSLKLANIVERYSSFLDALINRAARAPCOARDKVKLL
CLTDRALLCFDEPALLEHOVIGIDDAFLORELQDLQALQDSREHTEALQLL
KQDLAETKSS"KSLATTIGEFERLHRLIRKOKAMLELEADTARTLTDIBOKVRY
SQOLKVBQGAQILQERLAETDRHTFLAGVASLSERLKGKIHETLWTVDEPTSKVTG
PHYITWKSLODILHPFAALUDPGTAHQRLILSDDCITIVAYGNLHQPLQDSKPRF
DVEVSLGSEAFSSFWVHVWVVAEKTOMVIGLAHEAASRGSIOIQPSRGFYCIVMH
DGNQYSACTEPWRLNRDKLVGVFLDYDQGLLI FYNADMSWLYTFRKFPGLKLC
SYFSPQSHANGKNVQLRINTVLI"

```

ORIGIN

```

Query Match      82.7%; Score 3163; DB 6; Length 3436;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3433; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 370 CCACTACCTCTAGATGCCCTCCCGGGCTGGGTCCACGGAGTCTCAGCGCGCACCCC 429
DB 1 CCACTACCTCTAGATGCCCTCCCGGGCTGGGTCCACGGAGTCTCAGCGCGCACCCC 60

QY 430 TTCCTCGGCTTACCTTCCTCCGACAGCACCCCTCCCTTCTCCGCTAGCTCCTACCCC 489
DB 61 TTCCTCGGCTTACCTTCCTCCGACAGCACCCCTCCCTTCTCCGCTAGCTCCTACCCC 120

QY 490 TGCCTGTGGGGCTGTCTCCCGCGCCGACCCCTCGGTCTGCTCCGACAGCGCGCGC 549
DB 121 TGCCTGTGGGGCTGTCTCCCGCGCCGACCCCTCGGTCTGCTCCGACAGCGCGCGC 180

QY 550 TCTCTCAGCGCCCTCGCCCTCGGGCCCTCTCTCTGTCGCCCTCGCGCATGGCG 609
DB 181 TCTCTCAGCGCCCTCGCCCTCGGGCCCTCTCTCTGTCGCCCTCGCGCATGGCG 240

QY 610 TGCAGCTCAAGAGCAGCTGTGTCTCATCTGCTGAGCATCTACAGGACCCCGGTG 669
DB 241 TGCAGCTCAAGAGCAGCTGTGTCTCATCTGCTGAGCATCTACAGGACCCCGGTG 300

QY 670 AGCTGGGTGCGAGCACTACTTCTGCGCGCTGTGATCAGGAGCACTGGGTGGCGAG 729
DB 301 AGCTGGGTGCGAGCACTACTTCTGCGCGCTGTGATCAGGAGCACTGGGTGGCGAG 360

QY 730 GAGCGCAGGGCGCCCGCACTCCCGAGTGGCGGCACTGTCGCGAGCGCGCGCTG 789
DB 361 GAGCGCAGGGCGCCCGCACTCCCGAGTGGCGGCACTGTCGCGAGCGCGCGCTG 420

QY 790 GCGCCAGCCTCAAGCTGGCCAAACATCGTGGAGCGCTACAGCTCTCTCCGCTGGACGCC 849
DB 421 GCGCCAGCCTCAAGCTGGCCAAACATCGTGGAGCGCTACAGCTCTCTCCGCTGGACGCC 480

QY 850 ATCTCAACGCGCGCGCGCGGACCTCGGACCGGACCGACGACGAGTCAAGCTCTTC 909
DB 481 ATCTCAACGCGCGCGCGCGGACCTCGGACCGGACCGGACGAGTCAAGCTCTTC 540

QY 910 TGCTCAGGACCGCGCGCTTCTGTCTCTTCTGGGACGAGCTGCACTGCAACGAGCAG 969
DB 541 TGCTCAGGACCGCGCGCTTCTGTCTCTTCTGGGACGAGCTGCACTGCAACGAGCAG 600

QY 970 CATCAGGTCAACGCGCATCGACGCGCTTCGACGAGCTGACGAGGAGTGAAGGACCAA 1029
DB 601 CATCAGGTCAACGCGCATCGACGCGCTTCGACGAGCTGACGAGGAGTGAAGGACCAA 660

QY 1030 CTTTCAGGCTTCAAGACAGGAGCGGACACACGAGCGCTGCGAGCTGCTCAAGCGA 1089
DB 661 CTTTCAGGCTTCAAGACAGGAGCGGACACACGAGCGCTGCGAGCTGCTCAAGCGA 720

QY 1090 CAATGGCGAGACCAAGTCTTCCACCAAGAGCTGCGGACCACTATTCGCGGAGGCTTC 1149
DB 721 CAATGGCGAGACCAAGTCTTCCACCAAGAGCTGCGGACCACTATTCGCGGAGGCTTC 780

QY 1150 GAGCGGTGACCGGCTGTGCTGTGAACGCGAGAGGCCCATGTGTAGGAGCTGGAGCGG 1209

```

```

DB 781 GAGCGGCTGCACCGGCTGCTGGTGAACGCCAGAGAGGCCATGCTAGAGAGCTGGAGGCG 840
QY 1210 GACACGCGCCGACCGCTGACGACATCGAGCAGAAAGTCCAGCGTCTACAGCCAGCAGCTG 1269
DB 841 GACACGCGCCGACCGCTGACGACATCGAGCAGAAAGTCCAGCGTCTACAGCCAGCAGCTG 900
QY 1270 CGCAAGGTCCAGAGGAGGACCCAGATCTTCGAGGAGCGGCTGGTGAACCCAGCCGGAC 1329
DB 901 CGCAAGGTCCAGAGGAGGACCCAGATCTTCGAGGAGCGGCTGGTGAACCCAGCCGGAC 960
QY 1330 ACCTTCCTGGCTGGGGTGGCTCTCACTGTCGAGCGGCTCAAGGGAARATTCATGAGACC 1389
DB 961 ACCTTCCTGGCTGGGGTGGCTCTCACTGTCGAGCGGCTCAAGGGAARATTCATGAGACC 1020
QY 1390 AACCTCAATATGAAGACTTCCCGACCTCCAAAGTACACAGGCCCCCTGCAGTACACCATC 1449
DB 1021 AACCTCAATATGAAGACTTCCCGACCTCCAAAGTACACAGGCCCCCTGCAGTACACCATC 1080
QY 1450 TGGAGTCCCTGTTCCAGGACATCCACCCAGTGCAGCGCCCTTAACCTTGACCCCGGGC 1509
DB 1081 TGGAGTCCCTGTTCCAGGACATCCACCCAGTGCAGCGCCCTTAACCTTGACCCCGGGC 1140
QY 1510 ACAGCCACACGCGCTGATCTGTCGGACGACTGCACCATTTGGCTTACGSCAACTTG 1569
DB 1141 ACAGCCACACGCGCTGATCTGTCGGACGACTGCACCATTTGGCTTACGSCAACTTG 1200
QY 1570 CACCACAGCCACTGAGGACTCGCCAAAGCGCTTCGATGTGGAGGTGTGGTGTCTGGGT 1629
DB 1201 CACCACAGCCACTGAGGACTCGCCAAAGCGCTTCGATGTGGAGGTGTGGTGTCTGGGT 1260
QY 1630 TCTGAAGCCTT CAGTAGTGGGTCCACTACTGCGAGGTGGTGGCGGAGAGACCCAG 1689
DB 1261 TCTGAAGCCTT CAGTAGTGGGTCCACTACTGCGAGGTGGTGGCGGAGAGACCCAG 1320
QY 1690 TGGGTGATCGGCTGSCACAGNAGCCGCAAGCGCAGGCGAGCATCCAGATCCAGCCC 1749
DB 1321 TGGGTGATCGGCTGSCACAGNAGCCGCAAGCGCAGGCGAGCATCCAGATCCAGCCC 1380
QY 1750 AGCGCGGCTTCTACTGTCATCGTATGTCAGTGGCAACAGTACAGCGCTGCACGGAG 1809
DB 1381 AGCGCGGCTTCTACTGTCATCGTATGTCAGTGGCAACAGTACAGCGCTGCACGGAG 1440
QY 1810 CCGTGAACGCGGCTTAACTCCGGGCAAGCTTGAACAGGTGGGTGTCTTCTTGGACTAT 1869
DB 1441 CCGTGAACGCGGCTTAACTCCGGGCAAGCTTGAACAGGTGGGTGTCTTCTTGGACTAT 1500
QY 1870 GACCAAGCTTCTCTCATCTTCTACATGCTGATGACATGTCCTGGCTCTACACCTTCCGC 1929
DB 1501 GACCAAGCTTCTCTCATCTTCTACATGCTGATGACATGTCCTGGCTCTACACCTTCCGC 1560
QY 1930 GAGAACTTCCCTGGCAAGCTCTGCTCTTACTTTCAGCCCTGGCCAGAGCCACGCAATGGC 1989
DB 1561 GAGAACTTCCCTGGCAAGCTCTGCTCTTACTTTCAGCCCTGGCCAGAGCCACGCAATGGC 1620
QY 1990 AGAAAGCTTCAGCGCTGGGATCAACACCGTCCGATCTAGTCCAGGAGAGAGGAGACC 2049
DB 1621 AGAAAGCTTCAGCGCTGGGATCAACACCGTCCGATCTAGTCCAGGAGAGAGGAGACC 1680
QY 2050 ACAACCTCTGGACCACTGCCACCTGCAAGAGCCCTGCCAGGAAGATAGAGAGACTGG 2109
DB 1681 ACAACCTCTGGACCACTGCCACCTGCAAGAGCCCTGCCAGG-AGATAGAGAGACTGG 1739
QY 2110 ACTCCAGCCACCGTGGCCACTGGAGACCTCAGGCGAGTTGTTTACCTCCAGCCCTCCAG 2169
DB 1740 ACTCCAGCCACCGTGGCCACTGGAGACCTCAGGCGAGTTGTTTACCTCCAGCCCTCCAG 1799
QY 2170 TCTGTAATAGGAGTTGATTCCTTCTTCTTAACTCTTTCAGCATCGATGTTCTG 2229
DB 1800 TCTGTAATAGGAGTTGATTCCTTCTTCTTAACTCTTTCAGCATCGATGTTCTG 1859
QY 2230 TAGCTCTGACCTTGATAGGATACAGCTTTGATCCAGAGATGTGACATGGCTTCTCTCTCA 2289
DB 1860 TAGCTCTGACCTTGATAGGATACAGCTTTGATCCAGAGATGTGACATGGCTTCTCTCTCA 1919

```



2290 GGGCAACCCCTGCGCCAAACCTCATCCCATCTCTCAGGGGAGGGGACTACCTTCCAGT 2349  
Db GGGCAACCCCTGCGCCAAACCTCATCCCATCTCTCAGGGGAGGGGACTACCTTCCAGT 1979  
2350 GTCTCCCTCCAGCCAGCCCTGACCTCAGAAAGTGTACAGATGCGCCAGTAGTGTGGCAG 2409  
Db GTCTCCCTCCAGCCAGCCCTGACCTCAGAAAGTGTACAGATGCGCCAGTAGTGTGGCAG 2039  
2410 CCGAAAGACACACAGCACCCCTCTTATGTCCCATGCGCTTAAGACTTACCCCTGACCAAGC 2469  
Db CCGAAAGACACACAGCACCCCTCTTATGTCCCATGCGCTTAAGACTTACCCCTGACCAAGC 2099  
2470 TAGTGATGGGCCAATTACCCCTTGACCCAGTCCACAGTGTGTACAGTAGTACTCGTTC 2529  
Db TAGTGATGGGCCAATTACCCCTTGACCCAGTCCACAGTGTGTACAGTAGTACTCGTTC 2159  
2530 TAGGTTGCTGAGAGCCAACTCTCTGCGCCACCCACACCAAGAACTATATGGTTCCCT 2589  
Db TAGGTTGCTGAGAGCCAACTCTCTGCGCCACCCACACCAAGAACTATATGGTTCCCT 2219  
2590 ACTTCTCCCACTGATCTGCTGTGAGTGATGTGTGCGCTGTGGAAGGCACTTGGA 2649  
Db ACTTCTCCCACTGATCTGCTGTGAGTGATGTGTGCGCTGTGGAAGGCACTTGGA 2279  
2650 GTTGAGTCCACATATAGTCAATGTGCGCACCACTCTCTGCGCCACAGGCGGAGGACAG 2709  
Db GTTGAGTCCACATATAGTCAATGTGCGCACCACTCTCTGCGCCACAGGCGGAGGACAG 2339  
2680 GTTGAGTCCACATATAGTCAATGTGCGCACCACTCTCTGCGCCACAGGCGGAGGACAG 2769  
Db GTTGAGTCCACATATAGTCAATGTGCGCACCACTCTCTGCGCCACAGGCGGAGGACAG 2399  
2770 GCCTCCCTGGATGATCGAGTCCCTCAGTAGCTCTGAAACAAGTCCAGCCAACTCTTC 2829  
Db GCCTCCCTGGATGATCGAGTCCCTCAGTAGCTCTGAAACAAGTCCAGCCAACTCTTC 2459  
2830 AGCCAGGCTCTGTGACCTGAGTGTGAGGAGGCTTCCAGAGCAGTGTGTTGTAATTA 2889  
Db AGCCAGGCTCTGTGACCTGAGTGTGAGGAGGCTTCCAGAGCAGTGTGTTGTAATTA 2519  
2890 GGCACCAAGCACTGGAGGGGCTGTGGCTAGACCCCTGTGACATTTGGCATCTATCTC 2949  
Db GGCACCAAGCACTGGAGGGGCTGTGGCTAGACCCCTGTGACATTTGGCATCTATCTC 2579  
2950 AGTTAGGATCTGTGACAGAAACAAGAGCAGTGTGAGTGTGTTAATTAGACAAGAT 3009  
Db AGTTAGGATCTGTGACAGAAACAAGAGCAGTGTGAGTGTGTTAATTAGACAAGAT 2639  
3010 TTACTACTGGCCCTGTGCTGCAAAATTTGGAGAGCTGGAGAGCAGACTCTG 3069  
Db TTACTACTGGCCCTGTGCTGCAAAATTTGGAGAGCTGGAGAGCAGACTCTG 2699  
3070 CTGAATTTCCAGGAACCTCCAGCGCCAGATTCATCTGTCTGTTGTGACAGGAAGCTG 3129  
Db CTGAATTTCCAGGAACCTCCAGCGCCAGATTCATCTGTCTGTTGTGACAGGAAGCTG 2759  
3130 CCCCCTATGCGAGGAGCCATATGCCAGAAAGTGTGATGTCAGAACTAGGCTCCCTC 3189  
Db CCCCCTATGCGAGGAGCCATATGCCAGAAAGTGTGATGTCAGAACTAGGCTCCCTC 2819  
3190 TGCCACGCTCGTGCAGCCAAATAGATGCTGTAGGCTGCGCTCTCCCACTTCACTCA 3249  
Db TGCCACGCTCGTGCAGCCAAATAGATGCTGTAGGCTGCGCTCTCCCACTTCACTCA 2879  
3250 GTTCCCAATCTAAATTTTACAAGAGATTTCTGTTTGGGGAACTTAAGTCAGATCCAGA 3309  
Db GTTCCCAATCTAAATTTTACAAGAGATTTCTGTTTGGGGAACTTAAGTCAGATCCAGA 2939  
3310 ACCTTGGCTGCAAGGGAGTGTGGAAATGTCAATTTCCCTAGAGGAAGTGTAGGTTGGTG 3369  
Db ACCTTGGCTGCAAGGGAGTGTGGAAATGTCAATTTCCCTAGAGGAAGTGTAGGTTGGTG 2999

3370 GAGCAAGCCCACTCGCTTTTCTGCCCACAGCATCAATCTGGAAGAACTCGGAGAGG 3429  
Db GAGCAAGCCCACTCGCTTTTCTGCCCACAGCATCAATCTGGAAGAACTCGGAGAGG 3059  
3430 GTGGAGTCCACATCTAGGTTGTCTGCCCCCTTGGCTCTATCCCTGCCAGAGTGGAA 3489  
Db GTGGAGTCCACATCTAGGTTGTCTGCCCCCTTGGCTCTATCCCTGCCAGAGTGGAA 3119  
3490 CTGGAGAGTGGGCTGCAAGACTGAGCCTAAATGTCTCCCGGCTTGTGCTTTCTTCT 3549  
Db CTGGAGAGTGGGCTGCAAGACTGAGCCTAAATGTCTCCCGGCTTGTGCTTTCTTCT 3179  
3550 AGTCTCTGGGCTTAGATTTCTGCACTTTGGGTCTCTGACACAAACACCACTCCCAAGTAG 3609  
Db AGTCTCTGGGCTTAGATTTCTGCACTTTGGGTCTCTGACACAAACACCACTCCCAAGTAG 3239  
3610 CCGGAAGAGCTAAACACAGGGGTCTTAAATGGCTGCCCCGCCACCGGCTCCCT 3669  
Db CCGGAAGAGCTAAACACAGGGGTCTTAAATGGCTGCCCCGCCACCGGCTCCCT 3299  
3670 TGGGCAAAAGGAATTTGTGAGCCTTACCCAACTTCAACTACAGAACTCTGGGCCACCC 3729  
Db TGGGCAAAAGGAATTTGTGAGCCTTACCCAACTTCAACTACAGAACTCTGGGCCACCC 3359  
3730 CAGCAGTATTTTAAATTTAAATTTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 3789  
Db CAGCAGTATTTTAAATTTAAATTTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 3419  
3790 TTAAGTTACAGATGTC 3806  
Db TTAAGTTACAGATGTC 3436

## RESULT 3

BD156501 3436 bp DNA linear PAT 17-JAN-2003  
LOCUS Primer for synthesizing full-length cDNA and use thereof.

DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

BD156501.1 GI:27862259  
JP 2002191363-A/11344.  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
Ota, T., Isogai, T., Nishikawa, T., Hayashi, K., Saito, K., Yamamoto, J.,  
Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Otsuki, T.  
Primer for synthesizing full-length cDNA and use thereof  
Patent: JP 2002191363-A 11344 09-JUL-2002;  
HELIX RESEARCH INSTITUTE

COMMENT  
OS Homo sapiens (human)  
PN JP 2002191363-A/11344  
PD 09-JUL-2002

PF 28-JUL-2000 JP 2000280990  
PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU  
SAITO,

PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,  
PI KEIICHI NAGAI, TETSUJI OTSUKI

PC C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/PC  
10,

PC C12P21/02, C12Q1/68//C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC  
Primer for synthesizing full-length cDNA and use thereof FH Key

FT CDS Location/Qualifiers  
(235).. (1659).  
10, Location/Qualifiers  
1. 3436  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"

FEATURES  
source

ORIGIN  
Query Match 82.7%; Score 3163; DB 6; Length 3436;



```
QY 2530 TAGGTTGCCTGAGAGCAACCTCTCTGCGCACCCGCCACACAAAGAACTATATGTTCTCT 2589
Db 2160 TAGGTTGCCCTGAGAGCAACCTCTCTGCGCACCCGCCACACAAAGAACTATATGTTCTCT 2219
QY 2590 ACTTCTCCACTGATCTGCTGCTGATGATGATGCTGTGGCTGTGGAGGACACCTGGTA 2649
Db 2220 ACTTCTCCACTGATCTGCTGCTGATGATGATGCTGTGGCTGTGGAGGACACCTGGTA 2279
QY 2650 GTTGAGTCCACATATATAGTCAATGTGCCACCACTTCTGCTGCCACAGGCCGAGGACAG 2709
Db 2280 GTTGAGTCCACATATATAGTCAATGTGCCACCACTTCTGCTGCCACAGGCCGAGGACAG 2339
QY 2710 GGTGAGGCTATACCCAAAGCTGATCCAGAGCCATTTAGCTTAAAGCACTCCAGACAA 2769
Db 2340 GGTGAGGCTATACCCAAAGCTGATCCAGAGCCATTTAGCTTAAAGCACTCCAGACAA 2399
QY 2770 GCCTCCCTGGATGATCGAGTCCCCAGTAGCTCTGAACAAGAGTCCAGCAACCTCTTTC 2829
Db 2400 GCCTCCCTGGATGATCGAGTCCCCAGTAGCTCTGAACAAGAGTCCAGCAACCTCTTTC 2459
QY 2830 AGCCAGGCTCTGTGACCTGTAGGCTGAGAGGCTTCCAGAGCACTGTGTTGTAATTA 2889
Db 2460 AGCCAGGCTCTGTGACCTGTAGGCTGAGAGGCTTCCAGAGCACTGTGTTGTAATTA 2519
QY 2890 GGACCCAGCACTGGAGGGGCTGTGGCTAGACCCCTGTGACACTTGGCACTATCTCTC 2949
Db 2520 GGACCCAGCACTGGAGGGGCTGTGGCTAGACCCCTGTGACACTTGGCACTATCTCTC 2579
QY 2950 AGTTAGGATCTGTGCTGCAGAAACAGAGCCACTTGTAGCTGGTTTAAATTAGACAGGAT 3009
Db 2580 AGTTAGGATCTGTGCTGCAGAAACAGAGCCACTTGTAGCTGGTTTAAATTAGACAGGAT 2639
QY 3010 TTACTACCTGGCCCTGTGGCTGTGCAAAATTTGTTGGAAGAGCTGGAGAAGCACTCTG 3069
Db 2640 TTACTACCTGGCCCTGTGGCTGTGCAAAATTTGTTGGAAGAGCTGGAGAAGCACTCTG 2699
QY 3070 CTGAATTTCCAGGAATCCAGGCGCAGATTCATCATGCTCTGTGACCAAGAAAGCTG 3129
Db 2700 CTGAATTTCCAGGAATCCAGGCGCAGATTCATCATGCTCTGTGACCAAGAAAGCTG 2759
QY 3130 CCCCACCTGCGAGGAAGCACTATGCGAAGAGCTGCTGACTGCGAGAACTAGGCTCCCTC 3189
Db 2760 CCCCACCTGCGAGGAAGCACTATGCGAAGAGCTGCTGACTGCGAGAACTAGGCTCCCTC 2819
QY 3190 TGCCACGCTCCGTGCGAGCAATAGATGCTGAGGCTGCCCCCTCTCCCACTTCACTCA 3249
Db 2820 TGCCACGCTCCGTGCGAGCAATAGATGCTGAGGCTGCCCCCTCTCCCACTTCACTCA 2879
QY 3250 GTTCCCAATCTAAATTTTCAAGAGATCTGTTGGGGGAACTTAAGTCAGATCCAGA 3309
Db 2880 GTTCCCAATCTAAATTTTCAAGAGATCTGTTGGGGGAACTTAAGTCAGATCCAGA 2939
QY 3310 ACCTTGGCTGCAAGGAGTCTGGGAAATGCTATTTCCCTAGAGGAAGTTAGGGTGGGTG 3369
Db 2940 ACCTTGGCTGCAAGGAGTCTGGGAAATGCTATTTCCCTAGAGGAAGTTAGGGTGGGTG 2999
QY 3370 GAGCAAGGCCCACTCGCTTTTCTGCCACAGATCCAAATGTAAGAACTCGGGAGAGG 3429
Db 3000 GAGCAAGGCCCACTCGCTTTTCTGCCACAGATCCAAATGTAAGAACTCGGGAGAGG 3059
QY 3430 GTGGAGTCCACATCTAGGGTGTGCTGCCCCCTGGCTCTATCCCTGCCAGAGGTGGAA 3489
Db 3060 GTGGAGTCCACATCTAGGGTGTGCTGCCCCCTGGCTCTATCCCTGCCAGAGGTGGAA 3119
QY 3490 CTGGAGAGTGGCTGCAAGACTGAGCTAAATGCTCTCCCGGCTTGTACTTTCTTCTCT 3549
Db 3120 CTGGAGAGTGGCTGCAAGACTGAGCTAAATGCTCTCCCGGCTTGTACTTTCTTCTCT 3179
QY 3550 AGTCTCTGGGCTTAGATCTTGCACTTTGGGCTCTCTGACACAAACACACCACTCCCAAGTAG 3609
Db 3180 AGTCTCTGGGCTTAGATCTTGCACTTTGGGCTCTCTGACACAAACACACCACTCCCAAGTAG 3239
```

```
QY 3610 CCGAAGAGCTAAACACAGGGGGTTCTTTAAATGCTGCCCGCCACCCCGGCTCCTCT 3669
Db 3240 CCGNAGAGCTAAACACAGGGGGTTCTTTAAATGCTGCCCGCCACCCCGGCTCCTCT 3299
QY 3670 TGGGCAAAAGGAATGTGAGCCCTACCCCAACCTTCAACTACCAAGATCTGGGCCACCC 3729
Db 3300 TGGGCAAAAGGAATGTGAGCCCTACCCCAACCTTCAACTACCAAGATCTGGGCCACCC 3359
QY 3730 CAGCAGTATTTTATTTAAATGTTGCCCATTTTATGAGTTATGATCAATTTGTATTAA 3789
Db 3360 CAGCAGTATTTTATTTAAATGTTGCCCATTTTATGAGTTATGATCAATTTGTATTAA 3419
QY 3790 TTAAAGTTACAGATGTC 3806
Db 3420 TTAAAGTTACAGATGTC 3436

AK001621 3436 bp mRNA linear PRI 01-AUG-2002
Homo sapiens cDNA FLJ10759 fis, clone NT2RP3004617, weakly similar
to ZINC-BINDING PROTEIN A33.
ACCESSION AK001621
VERSION AK001621.1 GI:7022986
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Isogai,T., Ota,T., Hayashi,K., Sugiyasu,T., Otsuki,T., Suzuki,Y.,
Nishikawa,T., Nagai,K., Sugano,S., Ishibashi,T., Fujimori,K.,
Tanai,H., Kimata,M., Watanabe,M., Hiraoka,S., Ishii,S., Kawai,Y.,
Saito,K., Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagahara,K.,
Masuo,Y. and Kanehori,K.
NEDO human cDNA sequencing project
Unpublished
2 (bases 1 to 3436)
Isogai,T. and Otsuki,T.
Direct Submission
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source
1..3436
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NT2RP3004617"
/cell_line="NT2"
/cell_type="teratocarcinoma"
/clone_lib="NT2RP3"
/note="cloning vector: pME18SFL3-mRNA from NT2 neuronal
precursor cells after 2-weeks retinoic acid (RA)
induction."
235..1662
/note="unnamed protein product"
/codon_start=1
/db_xref="GI:7022987"
/translation="MACSLKDELCSICLSIYQDPVSLGCEHYFCRCRITHEWROEA
QGARQPCPCRTFAEPALAPSILKLANIVERYSSFPDLAINARRARPQCADHKVKLL
CLTDKALLCFPCDEPALHEQVGTGDDAFDELQRELKQLQALQDSEREHTEALQLL
KQDAETKSKSLRTTIGAEFRLHRLRERQKAMLELEADTARTLTDIFQKQVRY
SQLEKRVQEGAIQLQERLAETDRHTFLAGVASLSERLKGKIHETNLTVDFTSKYTG
PLQYTIWKSIFQDIHPVPAALTLDPCTAHQRLILSDDCITVAYGNLHPQLQDSPKRF
```

DVEVSLGSEAFSGVHYEVVVAEKTQWVIGLAHEAASRKSIIQPSRFGYCIVMH  
 DNGYSACTEPWRLNVRDLKDVGFELDYDQGLIIFYNADDSWLYTFREKPGKLC  
 SYSPGSHANGKNVQPLRINTVLI"

ORIGIN

Query Match		82.7%; Score 3163; DB 9; Length 3436;
Best local Similarity		99.9%; Pred. No. 0;
Matches 3433; Conservative		0; Mismatches 3; Indels 1; Gaps 1;
QY	370	CCACTACTCTAGACTGCCCTCCGGGCTGCGCTCCACGAGTCTCAGCGCGCACCCC 429
DB	1	CCACTACTCTAGACTGCCCTCCGGGCTGCGCTCCACGAGTCTCAGCGCGCACCCC 60
QY	430	TTCTCTGCGGTTACCTCTTCCGGACAGACCCCTCCCTTCTCGGTAGTCTCTACCCC 489
DB	61	TTCTCTGCGGTTACCTCTTCCGGACAGACCCCTCCCTTCTCGGTAGTCTCTACCCC 120
QY	490	TGCTGTGCGGGCTGTCTCCCGGGCCAGCCCTCGGTGTGCTCTCGACGCGCGCC 549
DB	121	TGCTGTGCGGGCTGTCTCCCGGGCCAGCCCTCGGTGTGCTCTCGACGCGCGCC 180
QY	550	TCCTCTAGCGCGCCCTCGCCCTCGGGCCCTCTCTGCTGCGCTGCGCGCATGGCG 609
DB	181	TCCTCTAGCGCGCCCTCGCCCTCGGGCCCTCTCTGCTGCGCTGCGCGCATGGCG 240
QY	610	TGCAGCTCAAGGACGAGTGTGTGTCTCCATCTGCTGAGCATCTACAGGACCCGGTG 669
DB	241	TGCAGCTCAAGGACGAGTGTGTGTCTCCATCTGCTGAGCATCTACAGGACCCGGTG 300
QY	670	AGCTTGGGCTGCGAGCACTACTTCTGCGCGCGTGTGATCAGGAGCATGCGTGGCGAG 729
DB	301	AGCTTGGGCTGCGAGCACTACTTCTGCGCGCGTGTGATCAGGAGCATGCGTGGCGAG 360
QY	730	GAGCGCAGGCGCGCGGACTGCGGAGTGTGCGGCGCACCTTCTGCGGAGCCCGGCTG 789
DB	361	GAGCGCAGGCGCGCGGACTGCGGAGTGTGCGGCGCACCTTCTGCGGAGCCCGGCTG 420
QY	790	GCGCCACGCTOAGCTGGSCAAATCGTGAGCGCTTACAGCTCTTCCGCTGGAGGCC 849
DB	421	GCGCCACGCTCAAGCTGGSCAAATCGTGAGCGCTTACAGCTCTTCCGCTGGAGGCC 480
QY	850	ATCTCTAAACGCGCGCGCGCGGAGCTTCTGCGCGCGTGTGATCAGGAGCATGCGTCTTC 909
DB	481	ATCTCTAAACGCGCGCGCGCGGAGCTTCTGCGCGCGTGTGATCAGGAGCATGCGTCTTC 540
QY	910	TGCTCTACGAGCGCGCGCTTCTCTGCTTCTTCTGCGAGAGCTTGCAGTGCAGAGCAG 969
DB	541	TGCTCTACGAGCGCGCGCTTCTCTGCTTCTTCTGCGAGAGCTTGCAGTGCAGAGCAG 600
QY	970	CATCAGGTCAACGGCATCGACGAGCTTTCAGAGGTGCGAGGAGCTGAAGGACCAA 1029
DB	601	CATCAGGTCAACGGCATCGACGAGCTTTCAGAGGTGCGAGGAGCTGAAGGACCAA 660
QY	1030	CTTCAGGCGCTTCAAGACAGCGAGCGGGAACACACGAGCGCTGAGCTGCTCAAGCGA 1089
DB	661	CTTCAGGCGCTTCAAGACAGCGAGCGGGAACACACGAGCGCTGAGCTGCTCAAGCGA 720
QY	1090	CAACTGGCGGAGCAAGTCTTCCACCAAGAGCTGCGGACCACTATCGCGGAGGCTTC 1149
DB	721	CAACTGGCGGAGCAAGTCTTCCACCAAGAGCTGCGGACCACTATCGCGGAGGCTTC 780
QY	1150	GAGCGGTGCAACGGTGTGTGTGAAACGCGAGAGCCATGTAGAGGAGCTGGAGGCG 1209
DB	781	GAGCGGTGCAACGGTGTGTGTGAAACGCGAGAGCCATGTAGAGGAGCTGGAGGCG 840
QY	1210	GACACGCGCGCAGCTGACGAGCATGAGCAGAAAGTCCAGCGCTACAGCAGAGCTG 1269
DB	841	GACACGCGCGCAGCTGACGAGCATGAGCAGAAAGTCCAGCGCTACAGCAGAGCTG 900
QY	1270	CGAAGGTCCAGGAGGAGCCAGATCTCTGAGGAGCGGCTGGCTGAAACCGACCGGCAC 1329
DB	901	CGAAGGTCCAGGAGGAGCCAGATCTCTGAGGAGCGGCTGGCTGAAACCGACCGGCAC 960

QY	1330	ACCTTCTGCTGGGTGGCTCCTACTGTGCGAGCGGCTCAAGGAAAAATCCATGAGACC 1389
DB	961	ACCTTCTGCTGGGTGGCTCCTACTGTGCGAGCGGCTCAAGGAAAAATCCATGAGACC 1020
QY	1390	AACCTCATATGAAGACTTCCCGAAGCTTCAAGTACACAGGCGCCCTGCGAGTACCAATC 1449
DB	1021	AACCTCATATGAAGACTTCCCGAAGCTTCAAGTACACAGGCGCCCTGCGAGTACCAATC 1080
QY	1450	TGGAAGTCCCTGTTCAGGACATCCACCCAGTCCAGCGCCCTTAAACCTTGGACCCGGGC 1509
DB	1081	TGGAAGTCCCTGTTCAGGACATCCACCCAGTCCAGCGCCCTTAAACCTTGGACCCGGGC 1140
QY	1510	ACAGCCACACAGCGCTGATCTGTGCGAGCACTGCAACCAATTGTGGCTTACGGCAACTTG 1569
DB	1141	ACAGCCACACAGCGCTGATCTGTGCGAGCACTGCAACCAATTGTGGCTTACGGCAACTTG 1200
QY	1570	CACCCACAGCACTGCGAGCACTGCGCAAGCGCTTCGATGTGAGGCTGTGCTGCTGGGT 1629
DB	1201	CACCCACAGCACTGCGAGCACTGCGCAAGCGCTTCGATGTGAGGCTGTGCTGCTGGGT 1260
QY	1630	TCTGAAGCCTTCACTAGTGTGCGTCCACTACTGTGAGGTGTGTGGCGGAGAGACCCAG 1689
DB	1261	TCTGAAGCCTTCACTAGTGTGCGTCCACTACTGTGAGGTGTGTGGCGGAGAGACCCAG 1320
QY	1690	TGGGTGATCGGGTGGCACACGAAAGCGCAAGCGGCGAGCATCCAGATCCAGCCC 1749
DB	1321	TGGGTGATCGGGTGGCACACGAAAGCGCGCAAGCGGCGAGCATCCAGATCCAGCCC 1380
QY	1750	AGCGCGGCTTCTACTGATGCTGATGACGATGCGCAACCAAGTACAGCGCTGCGAGGAG 1809
DB	1381	AGCGCGGCTTCTACTGATGCTGATGACGATGCGCAACCAAGTACAGCGCTGCGAGGAG 1440
QY	1810	CCCTGAGCGCGCTTAACTGCGGCAAGCTTGAAGGTGGGTGTCTTCTGGAGCTAT 1869
DB	1441	CCCTGAGCGCGCTTAACTGCGGCAAGCTTGAAGGTGGGTGTCTTCTGGAGCTAT 1500
QY	1870	GACCAAGGCTTGTCTATCTTCTACAATGTGATGACATGCTTGGCTTACACCTTCCGC 1929
DB	1501	GACCAAGGCTTGTCTATCTTCTACAATGTGATGACATGCTTGGCTTACACCTTCCGC 1560
QY	1930	GAGAGTTCCTTGGAGCTTCTCTTACTTACGCGCTGGCGAGCGACGCAATGGC 1989
DB	1561	GAGAGTTCCTTGGAGCTTCTCTTACTTACGCGCTGGCGAGCGACGCAATGGC 1620
QY	1990	AAGAACGTTTCAGCGCTGCGGATCAACCGCTCGCATCTAGTCCAGGCAAGAGGAGACC 2049
DB	1621	AAGAACGTTTCAGCGCTGCGGATCAACCGCTCTCATCTAGTCCAGGCAAGAGGAGACC 1680
QY	2050	ACAACCTCTCTGGGACCACTGCCACCTGCAAGAGCCCTTGGCCAGGAGATGAAGACCTGG 2109
DB	1681	ACAACCTCTCTGGGACCACTGCCACCTGCAAGAGCCCTTGGCCAGG-AGA TAGAAGACCTGG 1739
QY	2110	ACTCGAGCGCACCGTGGCGACCTGAGAGCTTGGTTCAGCTTCCAGCTTCCAGCTTCCAG 2169
DB	1740	ACTCGAGCGCACCGTGGCGACCTGAGAGCTTGGTTCAGCTTCCAGCTTCCAGCTTCCAG 1799
QY	2170	TCTGTAAGTGGAGGTTCATCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2229
DB	1800	TCTGTAAGTGGAGGTTCATCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1859
QY	2230	TAGCTCTGACCTTGATAGGATACAGCTTTGATCCAAAGGATGTGACATGGCTTCTCTCA 2289
DB	1860	TAGCTCTGACCTTGATAGGATACAGCTTTGATCCAAAGGATGTGACATGGCTTCTCTCA 1919
QY	2290	GGGACACCTTGGCGACCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2349
DB	1920	GGGACACCTTGGCGACCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1979
QY	2350	GTCTTCTTCCAGCGCGCTGACCTTCAAGAGGTGTGAGAGCATGCGAGTAGTGGAG 2409
DB	1980	GTCTTCTTCCAGCGCGCTGACCTTCAAGAGGTGTGAGAGCATGCGAGTAGTGGAG 2039
QY	2410	CCGAAAGACACACAGCACCTCTTATGTGCCATGGCCTTAAGACTTACCCCTGACCAAGC 2469







1831 CGGACAAAGCTTGACAAAGTGGGTGCTCTCTGACATATGACCAAGGCTTCTCATCTTC 1890  
 1261 CGGACAAAGCTTGACAAAGTGGGTGCTCTCTGACATATGACCAAGGCTTCTCATCTTC 1320  
 1891 TACAATGCTGATGATGCTCTGCTCTACACCTTCGCGGAGAAAGTTCCTCTGGCAAGCTC 1950  
 1321 TACAATGCTGATGATGCTCTGCTCTACACCTTCGCGGAGAAAGTTCCTCTGGCAAGCTC 1380  
 1951 TGCTCTTACTTTCAGCCCTCGGACAGGACGACCAATGGCAAGAAAGTTCAGCGCTGGGG 2010  
 1381 TGCTCTTACTTTCAGCCCTCGGACAGGACGACCAATGGCAAGAAAGTTCAGCGCTGGGG 1440  
 2011 ATCAACACCGCTCGCATCTAGTTCGAGGAGAGAGACCAAACTCTGCGGACCACTGC 2070  
 1441 ATCAACACCGCTCGCATCTAGTTCGAGGAGAGAGACCAAACTCTGCGGACCACTGC 1500  
 2071 CACCTGCAAGAGCCCTGCGGAGGAGATAGAAAGCTTGAACCTCCAGCCCAACCGTGGCCAC 2130  
 1501 CACCTGCAAGAGCCCTGCGGAGGAGATAGAAAGCTTGAACCTCCAGCCCAACCGTGGCCAC 1559  
 2131 TGGAGACCTCAGGCGAGTGTGTTAACCTCCAGCCTCCAGCTCTGTAAATGGAGGTTCAT 2190  
 1560 TGGAGACCTCAGGCGAGTGTGTTAACCTCCAGCCTCCAGCTCTGTAAATGGAGGTTCAT 1619  
 2191 TCCCTACTTCTAACTCTCTCCAGCATCGATGTTCTGTAGCTCTGACCTGTATAGGGA 2250  
 1620 TCCCTACTTCTAACTCTCTCCAGCATCGATGTTCTGTAGCTCTGACCTGTATAGGGA 1679  
 2251 TACAGCTTTGATCCAGGATGTGATGGCTTCTCTCAGGCGCAACCCCTGCCCAACCT 2310  
 1680 TACAGCTTTGATCCAGGATGTGATGGCTTCTCTCAGGCGCAACCCCTGCCCAACCT 1739  
 2311 CATCCCCATCTTCTCAGGCGCAGGGAATACCTTCAGGCTCTCCCTCAGGCGCAGCT 2370  
 1740 CATCCCCATCTTCTCAGGCGCAGGGAATACCTTCAGGCTCTCCCTCAGGCGCAGCT 1799  
 2371 GACCTCAGGAAGTGTGAGAGATGCGCAGTGTGGCAGCCGGAAGACACACAGCAC 2430  
 1800 GACCTCAGGAAGTGTGAGAGATGCGCAGTGTGGCAGCCGGAAGACACACAGCAC - CC 1858  
 2431 TCTTATGCTCCATGCTGCTAAGACTTACCTCTGACCAAGCTAGTGAGGCGCACTTACCT 2490  
 1859 TCTTATGCTCCATGCTGCTAAGACTTACCTCTGACCAAGCTAGTGAGGCGCACTTACCT 1918  
 2491 TGACCCCACTCCACAGTGTGTCACAGGTAGTA CTTGGTCTCAGGTTGCTGAGAGCCAC 2550  
 1919 TGACCCCACTCCACAGTGTGTCACAGGTAGTA CTTGGTCTCAGGTTGCTGAGAGCCAC 1978  
 2551 CTCTCTGCGACCCCGACACCAAGAACTATATGTTCTTACTTCTCCCACTGATCTGCTG 2610  
 1979 CTCTCTGCGACCCCGACACCAAGAACTATATGTTCTTACTTCTCCCACTGATCTGCTG 2038  
 2611 GTCAAGTATGATGCTGTGGCTGTGGAAGCACTGTGTAGTTGAGTCCACACATATAGT 2670  
 2039 GTCAAGTATGATGCTGTGGCTGTGGAAGCACTGTGTAGTTGAGTCCACACATATAGT 2098  
 2671 CATGTGCCACCACTTCTCTGCGCACAGGCGGAGGACAGGCTGAGGTTATACCAAGCT 2730  
 2099 CATGTGCCACCACTTCTCTGCGCACAGGCGGAGGACAGGCTGAGGTTATACCAAGCT 2158  
 2731 GATGAGAGCCCACTTATAGCTTAAAGCACTGACAGGACCAAGCTTCTGATGATCGAGGT 2790  
 2159 GATGAGAGCCCACTTATAGCTTAAAGCACTGACAGGACCAAGCTTCTGATGATCGAGGT 2218  
 2791 CCCAGTGTCTGAAACAGAGTCCAGCAACCTCTTCCAGCCAGGCTCTGTGACCTGC 2850  
 2219 CCCAGTGTCTGAAACAGAGTCCAGCAACCTCTTCCAGCCAGGCTCTGTGACCTGC 2278  
 2851 TAGGGTGCAGGAGCTTCCAGAGAGCTTGTGTAAATTAGGACCCAGCACTGGAGGGG 2910  
 2279 TAGGGTGCAGGAGCTTCCAGAGAGCTTGTGTAAATTAGGACCCAGCACTGGAGGGG 2338

2911 CTGTGGCTAGACCCCTTGTGACAGCTTGGCATCTATCTCAGTTAGGATCTCTGCTGCAGAA 2970  
 2339 CTGTGGCTAGACCCCTTGTGACAGCTTGGCATCTATCTCAGTTAGGATCTCTGCTGCAGAA 2398  
 2971 AACAGAGCCACTTGTAGCTGTTTAAATTAGACAAGGATTTACTACCTGGCCCTGTGGGG 3030  
 2399 AACAGAGCCACTTGTAGCTGTTTAAATTAGACAAGGATTTACTACCTGGCCCTGTGGGG 2458  
 3031 CTTCCAAAATTTGTTGGAGAGCTGAGAGAGCTCTGTGAAATTTTCCAGGAACTCCCA 3090  
 2459 CTTGCAAAAATTTGTTGGAGAGCTGAGAGAGCTCTGTGAAATTTTCCAGGAACTCCCA 2518  
 3091 GCGCCAGATTCAATCATGTCTCTGTGACCAAGAAAGCTGCCCCCATCTGCGAGGAAGCCAC 3150  
 2519 GCGCCAGATTCAATCATGTCTCTGTGACCAAGAAAGCTGCCCCCATCTGCGAGGAAGCCAC 2578  
 3151 TATGCCAGAAAGCTGCTGACTGCAAGACTAGGCTCCCTCTGCGACGCTCCGTCGACAGCA 3210  
 2579 TATGCCAGAAAGCTGCTGACTGCAAGACTAGGCTCCCTCTGCGACGCTCCGTCGACAGCA 2638  
 3211 ATAGATGCTCTGAGGCTGCCCTCTCCCACTTCACTCAGTTTCCCAATCTAAATTTTAA 3270  
 2639 ATAGATGCTCTGAGGCTGCCCTCTCCCACTTCACTCAGTTTCCCAATCTAAATTTTAA 2698  
 3271 CAAGAGATTCTGTTTGGGGAACTTAAAGTCCAGAACCTTGGCTGCAAGGAGTCT 3330  
 2699 CAAGAGATTCTGTTTGGGGAACTTAAAGTCCAGAACCTTGGCTGCAAGGAGTCT 2758  
 3331 GGGAAATGTCATTTCCCTAGAGGAAGTTAGGGTGGTGAGCAAGCCCACTCGCTTT 3390  
 2759 GGGAAATGTCATTTCCCTAGAGGAAGTTAGGGTGGTGAGCAAGCCCACTCGCTTT 2818  
 3391 TTCTGCCACAGCATCCAATCGTGAAGAACTCGGGAGAGGCTGGAGTCCACATCTAGGGTT 3450  
 2819 TTCTGCCACAGCATCCAATCGTGAAGAACTCGGGAGAGGCTGGAGTCCACATCTAGGGTT 2878  
 3451 GTCTGCCCCCTTGGCTCTATCCCTGCCAGAGGTGGGAAGTGGAGAGTGGGCTGCAAGA 3510  
 2879 GTCTGCCCCCTTGGCTCTATCCCTGCCAGAGGTGGGAAGTGGAGAGTGGGCTGCAAGA 2938  
 3511 CTGAGCTTAAATGTCCTCCCGGCTTGAATTTCTTCTAGTCTGGGCTAGATTTCTG 3570  
 2939 CTGAGCTTAAATGTCCTCCCGGCTTGAATTTCTTCTAGTCTGGGCTAGATTTCTG 2998  
 3571 CACTTGGGGTCTCTGACACACACCACTCCCAAGTAGTCCGGAAGAGCTTAAACACAGGG 3630  
 2999 CACTTGGGGTCTCTGACACACACCACTCCCAAGTAGTCCGGAAGAGCTTAAACACAGGG 3058  
 3631 GGTTCCTAAAATGGCTGCCCCGCGCACCCGGGCTCCCTTGGGCAAAAGAAATGTGAGC 3690  
 3059 GGTTCCTAAAATGGCTGCCCCGCGCACCCGGGCTCCCTTGGGCAAAAGAAATGTGAGC 3118  
 3691 CCTACCCGAACCTTCAACTACCAAGATCTGGGCCACCCAGCAGTATTTTATTTAAAA 3750  
 3119 CCTACCCGAACCTTCAACTACCAAGATCTGGGCCACCCAGCAGTATTTTATTTAAAA 3178  
 3751 TGTGTGCCATTTTATGAGTTATGATCAATTTGTTATTTAAATTTAAAGTTACAGATGTCAAAA 3810  
 3179 TGTGTGCCATTTTATGAGTTATGATCAATTTGTTATTTAAATTTAAAGTTACAGATGTCAAAA 3238  
 3811 AAAAAAATAAAAAA 3826  
 3239 AAAAAAATAAAAAA 3254

RESULT 6  
 BC007999  
 LOCUS  
 DEFINITION  
 Homo sapiens hypothetical protein FLJ.0759, mRNA (cDNA clone  
 MGC:15757 IMAGE:3357436), complete cds.  
 ACCESSION  
 BC007999  
 VERSION  
 BC007999.1  
 GI:14124949  
 KEYWORDS  
 MGC.



SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 3259)
AUTHORS	Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L.H., Shenmen,C.M., Schuler,G.D., Altshul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Haieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullay,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Wray,K.E., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Willalson,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettelman,M., Madan,A., Rodrigues,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butlerfield,Y.S., Krzywinski,M.I., Skalski,U., Smalys,D.E., Scherch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
MEDLINE	22388257
PUBMED	12477932
REFERENCE	2 (bases 1 to 3259)
AUTHORS	Strausberg,R.
TITLE	Direct Submission
JOURNAL	Submitted (15-MAY-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK	NIH-MGC Project URL: <a href="http://mgc.nci.nih.gov">http://mgc.nci.nih.gov</a>
COMMENT	Contact: MGC help desk Email: <a href="mailto:cgabps-remail.nih.gov">cgabps-remail.nih.gov</a> Tissue Procurement: ATCC cDNA Library Preparation: Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada <a href="mailto:info@bcgsc.bc.ca">info@bcgsc.bc.ca</a> Steven Chan, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Jones, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Ness, Fawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranada Teai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.
FEATURES	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <a href="http://image.llnl.gov">http://image.llnl.gov</a> Series: IRAL Plate: 23 Row: a Column: 20. Location/Qualifiers 1. 3259 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="MGC:15757 IMAGE:3357436" /tissue_type="Eye, retinoblastoma" /clone_lib="NIH MGC 16" /lab_host="DH10B-R" /note="Vector: pOTF7" 1. 3259 /gene="FLJ10759" /db_xref="LocusID:55223" 34. 1461 /codon_start=1 /product="Hypothetical protein FLJ10759" /protein_id="AAH07999.1"
gene	
CDS	



```

Db      2819  |||||TTCGCCACATCCAAATCGTGAAGAACTCGGAGAGGGTGGAGTCACATCTAGGGTT 2878
Qy      3451  GTCTGCGCCCTTGCTCTATCCCTGCGGAGGTGGGAATCGGAGGAGTGGCTGCAAGA 3510
Db      2879  |||||TTCGCCCTTGCTCTATCCCTGCGGAGGTGGGAATCGGAGGAGTGGCTGCAAGA 2938
Qy      3511  CTGAGCCTAAATGCTCTCCCGGCTTGACATTTCTTTCTAGTCTCTGGGCTAGATTCG 3570
Db      2939  CTGAGCCTAAATGCTCTCCCGGCTTGACATTTCTTTCTAGTCTCTGGGCTAGATTCG 2998
Qy      3571  CACTTGGGGTCTGACACACACACCATCCCAAGTAGCGGAGAGCTAAACACAGGG 3630
Db      2999  CACTTGGGGTCTGACACACACACCATCCCAAGTAGCGGAGAGCTAAACACAGGG 3058
Qy      3631  GGTTCTTAAATGCTCGCCCGGCTTGACATTTCTTTCTAGTCTCTGGGCTAGATTCG 3690
Db      3059  GGTTCTTAAATGCTCGCCCGGCTTGACATTTCTTTCTAGTCTCTGGGCTAGATTCG 3118
Qy      3691  CTTACCCCAACCTTCAATACCAAGATCTGGGCTAGATTTTATTTTAAAA 3750
Db      3119  CTTACCCCAACCTTCAATACCAAGATCTGGGCTAGATTTTATTTTAAAA 3178
Qy      3751  TGTGCGCATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3810
Db      3179  TGTGCGCATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3238
Qy      3811  AAAAAAAAAAAAAA 3826
Db      3239  AAAAAAAAAAAAAA 3254

RESULT 7
BC011689
LOCUS
DEFINITION
Homo sapiens hypothetical protein FLJ10759, mRNA (cDNA clone
MGC:19672 IMAGE:3353034), complete cds.
VERSION
BC011689
KEYWORDS
MGC.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 3259)
Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Ustin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullahy,S.J., Bosak,S.A., McMan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalusz,D.E.,
Scherer,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
2 (bases 1 to 3259)
Strausberg,R.
Direct Submission
TITLE
Submitted (30-JUL-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

```

```

REMARK
COMMENT
NIH-MGC Project URL: http://mgc.nci.nih.gov
On Dec 19, 2003 this sequence version replaced gi:15079757.
Contact: MGC help desk
Email: cgabps-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: http://www.nisc.nih.gov/
Contact: nisc_mgc@nhgri.nih.gov
Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R.,
Maduro,Q.L., Masello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C.,
McDowell,J., Pearson,R., Stantropop,S., Thomas,P.J., Touchman,J.W.,
Teurgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,
Young,A., Zhang,L.-H. and Green,E.D.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAL Plate: 27 Row: B Column: 5
This clone was selected for full length sequencing because it
passed the following selection criteria: Similarity but not
identity to protein.
Location/Qualifiers
1. .3259
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:19672 IMAGE:3353034"
/tissue_type="Eye, retinoblastoma"
/clone_lib="NIH_MGC_16"
/lab_host="DH10B-R"
/notes="Vector: pOTB7"
gene
1. .3259
/gene="FLJ10759"
/db_xref="LocusID:55223"
34. .1461
/gene="FLJ10759"
/codon_start=1
/product="hypothetical protein FLJ10759"
/protein_id="AAH11689.1"
/db_xref="GI:15079758"
/db_xref="LocusID:55223"
/translation="MACSLKDELLCSICLSIYQDPVSLGCEHYFCRCITEHWVRQEA
QGAARDCEPRTFAEPALAPSLKIANIVRYSSFPDLAINAPRAAPCOAHDKVILF
CLTDALLCFCDPEALHGHQVTHDDAPDELORELKQALQADSERHTEALQIL
KRLAETKSTSLRTTIGEAFERHLRLEROKAMLELEADTATLTDIECKVORY
SQOLRVQSGAQLQERLAETDHTFLAGVASLSERLKGKIHNTIYEDFPFSKVTG
FQYTTKSLFQDIHPVPAALTLDPGTAHQRLISDDCTIVAYGNLHPQLQDSPRRF
DGVSVLGSEAFSSGVHYVEVVAETQWVGLAHEAASRGKSIQIPRGFGYCIYVH
DGNYSACTEPTRLNVKDKLVGVFLDYDQGLLIYFNADDMWLTTFREKPPGKLC
SYFSPQCSHANGKNVQPLRINTVRI"
64. .204
/gene="FLJ10759"
/notes="RING; Region: RING-finger (Really Interesting New
Gene) domain, a specialized type of Zn-finger of 40 to 60
residues that binds two atoms of zinc"
/db_xref="CDD:cd00162"
445. .798
/gene="FLJ10759"
/notes="BEC; Region: B-Box C-terminal domain"
/db_xref="CDD:smart00502"
913. .1074
/gene="FLJ10759"
/notes="PRY; Region: associated with SPRY domains"
/db_xref="CDD:smart00589"
1075. .1413
/gene="FLJ10759"
/notes="SPRY; Region: Domain in SPLa and the RYanodine

```



Db 2039 GTCAATGATGATGCTGGGCTGTGGAAGCACTGTGATGTTGAGTCCACACATTAAGT 2098

QY 2671 CATGTGCCACCACTTCTCCGCCACAGCGCCAGGACAGAGGTGAGGGTATACCCAAAGCT 2730

Db 2099 CATGTGCCACCACTTCTCCGCCACAGCGCCAGGACAGAGGTGAGGGTATACCCAAAGCT 2158

QY 2731 GATGCAGAGCCCATTAAGCACTTAAAGCACTGCGAGCAAGCCCTCCCTGGATGATCGAGGT 2790

Db 2159 GATGCAGAGCCCATTAAGCACTTAAAGCACTGCGAGCAAGCCCTCCCTGGATGATCGAGGT 2218

QY 2791 CCCAGTGTAGCTCTGAACAGAGTCCAGCCAGCCCAACCTCTTACGACAGGCTCTGTGACCTGC 2850

Db 2219 CCCAGTGTAGCTCTGAACAGAGTCCAGCCCAACCTCTTACGACAGGCTCTGTGACCTGC 2278

QY 2851 TAGGGTGCAGAGGCTTCCAGAGCAGTGTCTGTGTAATTAGGACCCAGCACTGGGAGGG 2910

Db 2279 TAGGGTGCAGAGGCTTCCAGAGCAGTGTCTGTGTAATTAGGACCCAGCACTGGGAGGG 2338

QY 2911 CTGTTGGCTAGACCCCTTGTGCAGACTTGGCATCTATCTCAGTTAGGATCTGCTGCAGAA 2970

Db 2339 CTGTTGGCTAGACCCCTTGTGCAGACTTGGCATCTATCTCAGTTAGGATCTGCTGCAGAA 2398

QY 2971 AACAGAGCCACTGTAGCTGGTTTAATTAGACAAGGATTTACTACTGCGCCCTGGTGG 3030

Db 2399 AACAGAGCCACTGTAGCTGGTTTAATTAGACAAGGATTTACTACTGCGCCCTGGTGG 2458

QY 3031 CTTGCAAAATTTGTGGAAGAGCTGGAGAGCAGACTCTGCTGAATTTCCAGGAACCTCCCA 3090

Db 2459 CTTGCAAAATTTGTGGAAGAGCTGGAGAGCAGACTCTGCTGAATTTCCAGGAACCTCCCA 2518

QY 3091 GCGCCAGATTCATGCTGTGTTGTGACACGAGAAAGTGCCTCCCATCTGCGAAGCCAC 3150

Db 2519 GCGCCAGATTCATGCTGTGTTGTGACACGAGAAAGTGCCTCCCATCTGCGAAGCCAC 2578

QY 3151 TATGCAGAAAGCTGTGCTGCAGAACTAGGCTCCCTCTGCACCGTCCGTCGCGACCA 3210

Db 2579 TATGCAGAAAGCTGTGCTGCAGAACTAGGCTCCCTCTGCACCGTCCGTCGCGACCA 2638

QY 3211 ATAGATGTCTGTAGGGCTGCCCTCTCCCACTTCACTCACTGTTCCCAAACTCAATTTTAA 3270

Db 2639 ATAGATGTCTGTAGGGCTGCCCTCTCCCACTTCACTCACTGTTCCCAAACTCAATTTTAA 2698

QY 3271 CAAGAGATTCGTTTGGGGAACCTTAAGTCAGATCCAGAACCTTGGCTGCAAGGAGTCT 3330

Db 2699 CAAGAGATTCGTTTGGGGAACCTTAAGTCAGATCCAGAACCTTGGCTGCAAGGAGTCT 2758

QY 3331 GGGATATGTCATTTCCCTAGAGGAAGTTAGGTTGGTGGAGCAAGCCCACTGCTGTT 3390

Db 2759 GGGATATGTCATTTCCCTAGAGGAAGTTAGGTTGGTGGAGCAAGCCCACTGCTGTT 2818

QY 3391 TTCTGCCACAGCATCAATCGTGAAGAACTCGGGAGAGGGTGGAGTCCACATCTAGGGTT 3450

Db 2819 TTCTGCCACAGCATCAATCGTGAAGAACTCGGGAGAGGGTGGAGTCCACATCTAGGGTT 2878

QY 3451 GTCTGCCCCCTGGCTCTATCCCTGCCAGAGTGGGAACCTGAGAGTGGGCTGCAAGA 3510

Db 2879 GTCTGCCCCCTGGCTCTATCCCTGCCAGAGTGGGAACCTGAGAGTGGGCTGCAAGA 2938

QY 3511 CTGAGCTTAAATGCTCTCCCGGCTTGACTTTTCTTCTAGTCTCGGGCTAGATTCTG 3570

Db 2939 CTGAGCTTAAATGCTCTCCCGGCTTGACTTTTCTTCTAGTCTCGGGCTAGATTCTG 2998

QY 3571 CACTTGGGCTCTGTGACACACACCATCCCAAGTAGCCGGAAGAGCTAAACACAGGG 3630

Db 2999 CACTTGGGCTCTGTGACACACACCATCCCAAGTAGCCGGAAGAGCTAAACACAGGG 3058

QY 3631 GGTCTTAAATGGCTGCCCCGCCACCGGCTCCCTTGGCAGAAAGGAATGTCAGC 3690

Db 3059 GGTCTTAAATGGCTGCCCCGCCACCGGCTCCCTTGGCAGAAAGGAATGTCAGC 3118

QY 3691 CCTTACCCCAACCTTCACTACAGAACTCTGGGCCACCCAGCAGTATTTTATTTTAA 3750

Db 3119 CCTTACCCCAACCTTCACTACAGAACTCTGGGCCACCCAGCAGTATTTTATTTTAA 3178

QY 3751 TGTGCCCATTTTATGATGATCAATTTGTTATTAAATTAAGTTTACAGATGTCATAA 3810

Db 3179 TGTGCCCATTTTATGATGATGATCAATTTGTTATTAAATTAAGTTTACAGATGTCATAA 3238

QY 3811 AAAAAAAAAAAAAA 3826

Db 3239 AAAAAAAAAAAAAA 3254

RESULT 8  
BC012152 3259 bp mRNA linear PRI 04-OCT-2003  
LOCUS MGC:20370 IMAGE:4558639, complete cds.  
DEFINITION Homo sapiens hypothetical protein FLJ10759, mRNA (cDNA clone MGC:20370 IMAGE:4558639), complete cds.  
ACCESSION BC012152  
VERSION BC012152.1 GI:15082475  
KEYWORDS MGC.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 3259)  
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.P., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Scapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carrinci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullaly,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Hellon,E., Kettman,M., Madan,A., Rodriguez,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalusz,D.E., Scherch,A., Schein,J.E., Jones,S.J. and Marra,M.A.  
TITLE Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences  
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
PUBMED 22388257  
12477932  
2 (bases 1 to 3259)  
Strausberg,R.  
Direct Submission  
Submitted (02-AUG-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA  
NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgabbs-r@mail.nih.gov](mailto:cgabbs-r@mail.nih.gov)  
Tissue Procurement: ATCC  
CDNA Library Preparation: Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Genome Sequence Centre,  
BC Cancer Agency, Vancouver, BC, Canada  
[info@bcgsc.bc.ca](mailto:info@bcgsc.bc.ca)  
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline Schein, Duane Smalusz, Michael Smith, Lorraine Spence, Jeff Scott, Michael Thorne, Miranada Teai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAL Plate: 29 Row: d Column: 2  
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 8922647.  
Location/Qualifiers  
source  
1. .3259  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="MGC:20370 IMAGE:4558639"  
/tissue\_type="Eye, retinoblastoma"  
/clone\_lib="NIH MGC 16"  
/lab\_host="DH10B-R"  
/note="Vector: POTB7"  
1. .3259  
/gene="FLJ10759"  
/db\_xref="LocusID:55223"  
34. .1461  
/codon\_start=1  
/product="hypothetical protein FLJ10759"  
/protein\_id="AAH12152.1"  
/db\_xref="GI:15082476"  
/db\_xref="LocusID:55223"  
/translation="MACSLKDELICSLSIYQPVSLGCRHYFCRRICITHHWVROEA QGARDCEPRFAPALAPALQIIVERYSSPFLDALINARRAARPCQAHDKVLF CLDRALLCRPDPALHEQHVITIDDAFDELORELKQLOALQADSREHTEALQIL KRQLAETKSSYKSLRTTIGEPERHLRLREROKAMLEELADARTLTDIEQVORY SQQRKVOBQIILQERILAEFTDRHTFLAGVASLSERLKGKIHETNLTYEDPFTSKYTG PLYQTIWKSLEFQDILHPVPAALTDPGTAHQRLIISDDCTIIVAYGNLHPQPLQDSPKRF DGVSVLGSEAFSGVHWYVVAEKOTWIVGLAHEASRKSIOIOPSRGFYCIIVH DNQVYSACTEPWTRINVRDKLVGFLDYDQGLLIIFYNADMSWLYTFREKFPGLK SYFSPGSHANGKNVQPLRINTVRI"  
64. .192  
/note="RING; Region: Ring finger"  
/db\_xref="CDD:smart00184"  
445. .798  
/note="BBC; Region: B-Box C-terminal domain"  
/db\_xref="CDD:smart00502"  
913. .1074  
/note="SPRY; Region: associated with SPRY domains"  
/db\_xref="CDD:smart00589"  
1075. .1413  
/note="SPRY; Region: Domain in SPLa and the Ryanodine Receptor."  
/db\_xref="CDD:smart00449"  
ORIGIN  
Query Match 78.8%; Score 3014; DB 9; Length 3259;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 3254; Conservative 0; Mismatches 0; Indels 2; Gaps 2;  
QY 571 CCTGGGCCCCCTCTCTGCTGCCCTGGCGCCATGGGTGAGCCTCAAGGACGAGCTG 630  
Db 1 CCTGGGCCCCCTCTCTGCTGCCCTGGCGCCATGGGTGAGCCTCAAGGACGAGCTG 60  
QY 631 CTGTGCTCATCTGCTGAGCATCTACAGGACCCGGTGGCTGAGCTGGCTGAGCACTAC 690  
Db 61 CTGTGCTCATCTGCTGAGCATCTACAGGACCCGGTGGCTGAGCTGGCTGAGCACTAC 120  
QY 691 TTCTGCGCGCGCTGCATCAGGAGCACTGGGTGCGGCGAGGCGCGAGGCGCCGCGAC 750  
Db 121 TTCTGCGCGCGCTGCATCAGGAGCACTGGGTGCGGCGAGGCGCGAGGCGCCGCGAC 180  
QY 751 TGGCCGAGTGGCGGCGACGTTTGGCGAGCGCGCTGGCGCCACGCTCAAGCTGGCC 810  
Db 181 TGGCCGAGTGGCGGCGACGTTTGGCGAGCGCGCTGGCGCCACGCTCAAGCTGGCC 240  
QY 811 AACATGCTGAGCGCTACAGCTCTTCCCGTGGAGCCATCTCAACGCGCGCGCC 870  
Db 241 AACATGCTGAGCGCTACAGCTCTTCCCGTGGAGCGCCATCTCAACGCGCGCGCC 300  
QY 871 GCGGACCTTGGCGGCGACCAAGGTCAAGCTCTTCTGCTCTACGAGCGCGCGCTT 930  
Db 301 GCGGACCTTGGCGGCGACCAAGGTCAAGCTCTTCTGCTCTACGAGCGCGCGCTT 360



QY 2011 ATCAACACCGTCCGCACTATAGTCCAGGCAGAGGAGACCAACACCTCTCTGGGACCACTGC 2070  
Db 1441 ATCAACACCGTCCGCACTATAGTCCAGGCAGAGGAGACCAACCTCTCTGGGACCACTGC 1500  
QY 2071 CACCTGCAGAGCGCTGCCAGAGAGATAGAGACCTGGACTCGAGCCCAACCGTGGCCAC 2130  
Db 1501 CACCTGCAGAGCGCTGCCAGAGAGATAGAGACCTGGACTCGAGCCCAACCGTGGCCAC 1559  
QY 2131 TGGAGACCTCAGGCGCAGTGTGTTTACCTCCAGCTCCAGCTCTGTAAATGGAGTTGCAT 2190  
Db 1560 TGGAGACCTCAGGCGCAGTGTGTTTACCTCCAGCTCCAGCTCTGTAAATGGAGTTGCAT 1619  
QY 2191 TCCTACTTCTTAACTCTCTCCAGCATCGATGTTCTGTAGCTCTGACCTTGTATAGGGA 2250  
Db 1620 TCCTACTTCTTAACTCTCTCCAGCATCGATGTTCTGTAGCTCTGACCTTGTATAGGGA 1679  
QY 2251 TACAGCTTTGATCCAGGATGTGACATGGCTTCTCTCAGGGAACCCCTGCCCAACCT 2310  
Db 1680 TACAGCTTTGATCCAGGATGTGACATGGCTTCTCTCAGGGAACCCCTGCCCAACCT 1739  
QY 2311 CATCCCTCATCTTCTCAGGCGCAGGGACTACTCTTCCAGTGTCTCTCCCTCCAGCCGAGCCCT 2370  
Db 1740 CATCCCTCATCTTCTCAGGCGCAGGGACTACTCTTCCAGTGTCTCTCCCTCCAGCCGAGCCCT 1799  
QY 2371 GACCTCAGGAAGTGTGAGACATGGCGAGTAGTGTGGCAGCCCGAAAGACACACAGCACCC 2430  
Db 1800 GACCTCAGGAAGTGTGAGACATGGCGAGTAGTGTGGCAGCCCGAAAGACACACAGCACCC 1858  
QY 2431 TCTTATGTCCCATGGCTTAAGACTTACCTCTGACCAAGCTAGTGTAGTGGCCATTTACCT 2490  
Db 1859 TCTTATGTCCCATGGCTTAAGACTTACCTCTGACCAAGCTAGTGTAGTGGCCATTTACCT 1918  
QY 2491 TGAACCCAGTCCACAGTGTGCACAGTAGTACCTGTCTTAGGTTGCTCAGAGGCAAC 2550  
Db 1919 TGAACCCAGTCCACAGTGTGCACAGTAGTACCTGTCTTAGGTTGCTCAGAGGCAAC 1978  
QY 2551 CTCTCTGTCACCCCGCCACACCAAGAACTATATGTTCTCTTCTCCACTGTCTGCTG 2610  
Db 1979 CTCTCTGTCACCCCGCCACACCAAGAACTATATGTTCTCTTCTCCACTGTCTGCTG 2038  
QY 2611 GTCAAGTATGATGCTGTGGCTGTGGAAGCAGCTGTAGTGTAGTGTAGTGTAGT 2670  
Db 2039 GTCAAGTATGATGCTGTGGCTGTGGAAGCAGCTGTAGTGTAGTGTAGTGTAGT 2098  
QY 2671 CATGTGCCACACCTTCTGTCACAGGCGCGAGGACAGGTCAGGTCATACCCAAAGCT 2730  
Db 2099 CATGTGCCACACCTTCTGTCACAGGCGCGAGGACAGGTCAGGTCATACCCAAAGCT 2158  
QY 2731 GATGCAGAGGCCATTAGCTTAAAGCAACTGCAAGCAAGCTCTCTGTATGATGATGAGGT 2790  
Db 2159 GATGCAGAGGCCATTAGCTTAAAGCAACTGCAAGCAAGCTCTCTGTATGATGATGAGGT 2218  
QY 2791 CCCAGTGTCTGAACAGAGTCCAGGCAACCTCTTCCAGGCGGCTCTGTGACCTGC 2850  
Db 2219 CCCAGTGTCTGAACAGAGTCCAGGCAACCTCTTCCAGGCGGCTCTGTGACCTGC 2278  
QY 2851 TAGGGTGCAGGAGCTTCCAGAGCAGTGTGTTGTAATAGGACCCAGCACTGGAGGGG 2910  
Db 2279 TAGGGTGCAGGAGCTTCCAGAGCAGTGTGTTGTAATAGGACCCAGCACTGGAGGGG 2338  
QY 2911 CTGTTGGCTAGACCCCTGTGACACTTGGATCTATCTCAGTTAGATCTCTGTGAGAA 2970  
Db 2339 CTGTTGGCTAGACCCCTGTGACACTTGGATCTATCTCAGTTAGATCTCTGTGAGAA 2398  
QY 2971 AACAGAGCCACTGTAGCTGTTTAAATAGACAAGATTTACTACTGCGCCCTGGTGG 3030  
Db 2399 AACAGAGCCACTGTAGCTGTTTAAATAGACAAGATTTACTACTGCGCCCTGGTGG 2458  
QY 3031 CTGTGCAAAATTTGTGAAGAGCTTGAAGAGCAGACTCTGTAATTTCCAGGAACCTCCA 3090  
Db 2459 CTGTGCAAAATTTGTGAAGAGCTTGAAGAGCAGACTCTGTAATTTCCAGGAACCTCCA 2518  
QY 3091 GCGCCAGATTCATCATGTCTGTTGTGACCAAGAAAGCTGCCCTCATCTGCAGGAAGCCAC 3150

Db 2519 GCCCAGATTCATCATGTCTGTGTGACAGGAAGCTGCCCATCTGCAGGAAGCCAC 2578  
QY 3151 TATGCCAGAAAGCTGCTGACTGCAGAACTAGGTCCTCTGCAACGGTCCGTGCCAGCCA 3210  
Db 2579 TATGCCAGAAAGCTGCTGACTGCAGAACTAGGTCCTCTGCAACGGTCCGTGCCAGCCA 2638  
QY 3211 ATAGATGTCTCAGGCGCTGCCCTCTCCCACTTTCACCTCAGTCTCCCAATCTAAATTTT 3270  
Db 2639 ATAGATGTCTCAGGCGCTGCCCTCTCCCACTTTCACCTCAGTCTCCCAATCTAAATTTT 2698  
QY 3271 CAAGAGATTTCTGTTTGGGGAACCTTAAGTCAGATCCAGAACCTTGGCTGCAAGGAGTCT 3330  
Db 2699 CAAGAGATTTCTGTTTGGGGAACCTTAAGTCAGATCCAGAACCTTGGCTGCAAGGAGTCT 2758  
QY 3331 GGGAAATGTCTTCCCTAGAGGAAGTTAGGTTGGTGGTGGAGCAAGCCCACTGCGTTT 3390  
Db 2759 GGGAAATGTCTTCCCTAGAGGAAGTTAGGTTGGTGGTGGAGCAAGCCCACTGCGTTT 2818  
QY 3391 TTCTGCCCAGACATCCAAATCGTGAAGAACTCCGGAGAGGTTGAGTCCACATCTAGGTT 3450  
Db 2819 TTCTGCCCAGACATCCAAATCGTGAAGAACTCCGGAGAGGTTGAGTCCACATCTAGGTT 2878  
QY 3451 GTCTGCCCCCTTGGCTCTATCTCCCTGCCAGAGTGGGAACTGGAGAGTGGGCTGCAAGA 3510  
Db 2879 GTCTGCCCCCTTGGCTCTATCTCCCTGCCAGAGTGGGAACTGGAGAGTGGGCTGCAAGA 2938  
QY 3511 CTGAGCCTTAAATGTCTCCCGGCTTGAATTTCTTCTTAGTCTCTGGGCTAGATTTCTG 3570  
Db 2939 CTGAGCCTTAAATGTCTCCCGGCTTGAATTTCTTCTTAGTCTCTGGGCTAGATTTCTG 2998  
QY 3571 CACTTGGGCTCTCTGACACACACACATCCCAAGTGTAGCCGGAAGAGCTTAAACACAGGG 3630  
Db 2999 CACTTGGGCTCTCTGACACACACACATCCCAAGTGTAGCCGGAAGAGCTTAAACACAGGG 3058  
QY 3631 GGTCTTAAATGTCTCCCGGCTCCCGGCTCCCTTGGGCAAAAGGAATTTGTCAGC 3690  
Db 3059 GGTCTTAAATGTCTCCCGGCTCCCGGCTCCCTTGGGCAAAAGGAATTTGTCAGC 3118  
QY 3691 CTTACCCCAACCTTCAACTACAGAACTGGGCAACCCCGAGTATTTTATTTTAAAA 3750  
Db 3119 CTTACCCCAACCTTCAACTACAGAACTGGGCAACCCCGAGTATTTTATTTTAAAA 3178  
QY 3751 TGTGCCCCATTTTATGATTTATGATCAATTTGTATTAATTAAGTTACAGATGTCAAAA 3810  
Db 3179 TGTGCCCCATTTTATGATTTATGATCAATTTGTATTAATTAAGTTACAGATGTCAAAA 3238  
QY 3811 AAAAAAAAAAAAAA 3826  
Db 3239 AAAAAAAAAAAAAA 3254

RESULT 9  
AX775831

LOCUS AX775831 3243 bp mRNA linear PAT 14-JUL-2003  
DEFINITION Sequence 101 from Patent WO03048202.

ACCESSION AX775831  
VERSION AX775831.1 GI:32693549

KEYWORDS  
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1  
AUTHORS Matsuda, A. and Muramatsu, S.

TITLE NF-KB activating gene

JOURNAL Patent: WO 03048202-A 101 12-JUN-2003;  
Asahi Kasei Kabushiki Kaisha (JP)

FEATURES  
source  
1. .3243  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"



CDS	43..1470 /note= "unnamed protein product" /codon_start=1 /protein_id="CAE11615.1" /db_xref="GI:32693550" /translation="MACSKDELICSLISYQDPVSLGCEHYFCRCITEHWROEA GQARDPECFRTFAEPALPSLIANIVERYSFPLDALINARRARPQOAHDKVKLF CLTDRLCFFDEPFLHBEHQVTGIDDAFDLQKQQLQALQDSREHEALQLL KROLAETKSFKSTLTITGAERHLREROKAMELEADATRLTLDIEQKVRY SQOLAKYQEGAIIQERLAEDRHTFLAGVASLERLKIHETNLTIEDFPFTKYTG PLQYTIWKSJLPQDHPVPAAITLDPGTAHQRLILSDDTIVAYCNLHPQLDSPRF DEVSVLGSEAFSGVHIWEVVVAEKTKWVGIDLVQGLLIPYNADDMSWLTYTFREKPFGLK DNQYSACTEPTWELNRDLKDKGVGFELDYGILLIPYNADDMSWLTYTFREKPFGLK SYFSFGSHANGKNVQLFRINTVR1"
ORIGIN	
Query Match	78.3%; Score 2994; DB 6; Length 3243;
Best Local Similarity	99.9%; Pred. No. 0;
Matches 3234; Conservative	0; Mismatches 0; Indels 2; Gaps 2;
QY	571 CQTGGGCCCCCTCTGTGTCGCCCTGGCGCCCATGGCGTACAGCCTCAAGGACGAGCTG 630
DB	10 CTTGGGCCCCCTCTGTGTCGCCCTGGCGCCCATGGCGTACAGCCTCAAGGACGAGCTG 69
QY	631 CTGTGCTCCATCTGCCTGAGCATCTACAGAACCGCGTAGCGCTGGGGTGGAGCACTAC 690
DB	70 CTGTGCTCCATCTGCCTGAGCATCTACCAGAACCGCGTAGCGCTGGGGTGGAGCACTAC 129
QY	691 TTCTGCGCGCGCTGCATCACGGAGCACTGGGTGGCGGAGGAGCGCAGGGCGCCGCGAC 750
DB	130 TTCTGCGCGCGCTGCATCACGGAGCACTGGGTGGCGGAGGAGCGCAGGGCGCCGCGAC 189
QY	751 TGCCCGCGAGTGGCGCGCACGTTGCGCGAGCCCGCGCTGGCGCCAGCGCTCAAGCTGGCC 810
DB	190 TGCCCGCGAGTGGCGCGCACGTTGCGCGAGCCCGCGCTGGCGCCAGCGCTCAAGCTGGCC 249
QY	811 AACATCGTGGAGCGCTACAGTCTCTTCGCGTGGAGCGCCATCTCAAACGCGCGCGCGGCC 870
DB	250 AACATCGTGGAGCGCTACAGTCTCTTCGCGTGGAGCGCCATCTCAAACGCGCGCGCGGCC 309
QY	871 GGCGAACCTTGCAGCGGCGACAGCAAGGTCAAGTCTTCTGCGCTCAGGACCGCGCGCTT 930
DB	310 GGCGAACCTTGCAGCGGCGACAGCAAGGTCAAGTCTTCTGCGCTCAGGACCGCGCGCTT 369
QY	931 CTCTGCTTCTTTCGCGAGAGCGCTGCATCTGCACAGCAGCATCAGGTCAACCGGATCGAC 990
DB	370 CTCTGCTTCTTTCGCGAGAGCGCTGCATCTGCACAGCAGCATCAGGTCAACCGGATCGAC 429
QY	991 GACGCTTCGACGAGCTGCAGAGGAGCTGAAGACCAACTTCAGGCGGCTTCAAGACAGC 1050
DB	430 GACGCTTCGACGAGCTGCAGAGGAGCTGAAGACCAACTTCAGGCGGCTTCAAGACAGC 489
QY	1051 GAGCGGACACACCGAGCGCTGCAGTCTCAGCGCACCACTGGCGGAGACAAGTCT 1110
DB	490 GAGCGGACACACCGAGCGCTGCAGTCTCAGCGCACCACTGGCGGAGACAAGTCT 549
QY	1111 TCCACCAAGAGCTGGGACCACTATCGCGAGGCGCTTCGAGCGGCTGCACCGGCTGCTG 1170
DB	550 TCACCAAGAGCTGGGACCACTATCGCGAGGCGCTTCGAGCGGCTGCACCGGCTGCTG 609
QY	1171 CGTGAACGCGAAGGCCATGTAGAGGAGCTGGAGGCGGACACGCGCCGACCGTGAAC 1230
DB	610 CGTGAACGCGAAGGCCATGTAGAGGAGCTGGAGGCGGACACGCGCCGACCGTGAAC 669
QY	1231 GACATCGAGAGAAAGTCCAGCGCTACAGCGAGAGCTGGCGAAGTCCAGGAGGAGGCC 1290
DB	670 GACATCGAGAGAAAGTCCAGCGCTACAGCGAGAGCTGGCGAAGTCCAGGAGGAGGCC 729
QY	1291 CAGATCTGCAGGAGCGGCTGGCTGAAAACCGACCGGACACCTTCTGGCTGGGTGGCC 1350
DB	730 CAGATCTGCAGGAGCGGCTGGCTGMAACCGACCGGACACCTTCTGGCTGGGTGGCC 789
QY	1351 TCACTGTCCGAGCGCTCAAGGGAAAAATCCATATGAGACCAACCTCACATATGAAAGCTTC 1410
QY	



QY 61 CCAGGTTTGGGCGGGATCCGGCAGCTGAGCGGCGGCAACCCCTCTCTCTGCG 120  
 Db |||||  
 QY 76 CCAGGTTTGGGCGGGATCCGGCAGCTGAGCGGCGGCAACCCCTCTCTCTGCG 135  
 Db |||||  
 QY 121 GGTCAAGCAATGATGACGCTCGGCTGCTGCTCCCTCCCGCAGGATTCCTCCCA 180  
 Db |||||  
 QY 136 GGTCAAGCAATGATGACGCTCGGCTGCTGCTCCCTCCCGCAGGATTCCTCCCA 195  
 Db |||||  
 QY 181 GCTTCTCGGCTCCCGCAGCGCCACCGCCCGGATTCGACCCCTTAAAGGCTCCAC 240  
 Db |||||  
 QY 196 GCTTCTCGGCTCCCGCAGCGCCACCGCCCGGATTCGACCCCTTAAAGGCTCCAC 255  
 Db |||||  
 QY 241 CCGGCTCCGGGATCCCTTCTCCAGCTCTATCCCTTAGGACTGCCCGCCCTTAGAA 300  
 Db |||||  
 QY 256 CCGGCTCCGGGATCCCTTCTCCAGCTCTATCCCTTAGGACTGCCCGCCCTTAGAA 315  
 Db |||||  
 QY 301 CCTCCCGTCAAGATCTCGTCCCTGAGCGCTCAAGCTCTCTCCAGCGCCATGCGC 360  
 Db |||||  
 QY 316 CCTCCCGTCAAGATCTCGTCCCTGAGCGCTCAAGCTCTCTCCAGCGCCATGCGC 375  
 Db |||||  
 QY 361 TTGAGCTGCCATACCTTAGACTGCCCTCCGGGCTGGCTGCCACGGAGTCTCAGCC 420  
 Db |||||  
 QY 376 TTGAGCTGCCATACCTTAGACTGCCCTCCGGGCTGGCTGCCACGGAGTCTCAGCC 435  
 Db |||||  
 QY 421 GCGCAACCCCTTCTCGCGTTACCTCTTCCGGCAGCAGCCCTCCCTTCTCCGCTAGC 480  
 Db |||||  
 QY 436 GCGCAACCCCTTCTCGCGTTACCTCTTCCGGCAGCAGCCCTCCCTTCTCCGCTAGC 495  
 Db |||||  
 QY 481 TCTTACCCCTGCTGTGGGGCTCGTCCCGCGCCAGCCCTCGGTGTGCTCCGACA 540  
 Db |||||  
 QY 496 TCTTACCCCTGCTGTGGGGCTCGTCCCGCGCCAGCCCTCGGTGTGCTCCGACA 555  
 Db |||||  
 QY 541 GCGCGCGCTCTCTCAGCGCGCCCTGCGCTCCGGGCTGGCTGCCACGGAGTCTCAGCC 600  
 Db |||||  
 QY 556 GCGCGCGCTCTCTCAGCGCGCCCTGCGCTCCGGGCTGGCTGCCACGGAGTCTCAGCC 615  
 Db |||||  
 QY 601 GCGATGGGTGCGAGCTCAAGGACGAGCTGTGCTCCATCTGCTGAGCATCTACCA 660  
 Db |||||  
 QY 616 GCGATGGGTGCGAGCTCAAGGACGAGCTGTGCTCCATCTGCTGAGCATCTACCA 675  
 Db |||||  
 QY 661 GACCGGTGAGCTGGCTGCGAGCACTACTTCTGCGCGCTGATCAACGAGCACTGG 720  
 Db |||||  
 QY 676 GACCGGTGAGCTGGCTGCGAGCACTACTTCTGCGCGCTGATCAACGAGCACTGG 735  
 Db |||||  
 QY 721 GTGCGGAGGAGCGCAGCGCGCCCGGCTGCGCGAGTGGCGGCTGCGCGAG 780  
 Db |||||  
 QY 736 GTGCGGAGGAGCGCAGCGCGCCCGGCTGCGCGAGTGGCGGCTGCGCGAG 795  
 Db |||||  
 QY 781 CCGCGCTGGGCGCCAGCTCAAGCTGGCCAACTCGTGGAGCGCTACAGCTCTCTCCG 840  
 Db |||||  
 QY 796 CCGCGCTGGGCGCCAGCTCAAGCTGGCCAACTCGTGGAGCGCTACAGCTCTCTCCG 855  
 Db |||||  
 QY 841 CTGAGCGCATCTCAACGCGCGCGCGCGCGACCTGCGAGCGCAGCAAGGTC 900  
 Db |||||  
 QY 856 CTGAGCGCATCTCAACGCGCGCGCGCGCGACCTGCGAGCGCAGCAAGGTC 915  
 Db |||||  
 QY 901 AAGCTTCTGCTCTCAGGACCGCGGCTTCTCTGCTTCTTCTGCGAGCGCTGCACTG 960  
 Db |||||  
 QY 916 AAGCTTCTGCTCTCAGGACCGCGGCTTCTCTGCTTCTTCTGCGAGCGCTGCACTG 975  
 Db |||||  
 QY 961 CACGAGCAGCATCAGGTCAACGGCATCGACGAGCGCTTCTGAGAGCTGCGAGGGAGCTG 1020  
 Db |||||  
 QY 976 CACGAGCAGCATCAGGTCAACGGCATCGACGAGCGCTTCTGAGAGCTGCGAGGGAGCTG 1035  
 Db |||||  
 QY 1021 AAGGACCAATCTTCCGCGCTTCAAGCAGCGAGCGGGAACAACCGAAGCGCTGAGCTG 1080  
 Db |||||  
 QY 1036 AAGGACCAATCTTCCGCGCTTCAAGCAGCGAGCGGGAACAACCGAAGCGCTGAGCTG 1095  
 Db |||||  
 QY 1081 CTCAGGACCAATCTTCCGCGCTTCAAGCAGCGAGCGGGAACAACCGAAGCGCTGAGCTG 1140  
 Db |||||  
 QY 1096 CTCAGGACCAATCTTCCGCGCTTCAAGCAGCGAGCGGGAACAACCGAAGCGCTGAGCTG 1155  
 Db |||||

QY 1141 GAGGCTTTCAGCGGCTGCACCGGCTGCTGCTGAACGCGCAGAGGCACTGTAGAGGAG 1200  
 Db |||||  
 QY 1156 GAGGCTTTCAGCGGCTGCACCGGCTGCTGCTGAACGCGCAGAGGCACTGTAGAGGAG 1215  
 Db |||||  
 QY 1201 CTGAGGCGGACACCGGCGCGCACGCTGACCGCATCGAGCAAAAGTCCAGCGCTACAGC 1260  
 Db |||||  
 QY 1216 CTGAGGCGGACACCGGCGCGCACGCTGACCGCATCGAGCAAAAGTCCAGCGCTACAGC 1275  
 Db |||||  
 QY 1261 CAGCAGCTGCGCAAGGTCAGGAGGAGCCAGATCTCTGAGGAGCGGCTGCTGAACCC 1320  
 Db |||||  
 QY 1276 CAGCAGCTGCGCAAGGTCAGGAGGAGCCAGATCTCTGAGGAGCGGCTGCTGAACCC 1335  
 Db |||||  
 QY 1321 GACCGGCGCACCTTCTCTGCTGGGCTGCTCCTCTGCTGAGCGGCTCAAGGAAATAATC 1380  
 Db |||||  
 QY 1336 GACCGGCGCACCTTCTCTGCTGGGCTGCTCCTCTGCTGAGCGGCTCAAGGAAATAATC 1395  
 Db |||||  
 QY 1381 CATGAGCAAACTCATATATGAAGCTTCCGACCTCTCAAGTACACAGGCCCCCTGCGAG 1440  
 Db |||||  
 QY 1396 CATGAGCAAACTCATATATGAAGCTTCCGACCTCTCAAGTACACAGGCCCCCTGCGAG 1455  
 Db |||||  
 QY 1441 TACACCATCTGGAAGTCCCTGTTCCAGGACATCCACCGAGTCCAGCGCCCTTAACCTG 1500  
 Db |||||  
 QY 1456 TACACCATCTGGAAGTCCCTGTTCCAGGACATCCACCGAGTCCAGCGCCCTTAACCTG 1515  
 Db |||||  
 QY 1501 GACCGGCGCACAGCCACACAGCGCTGATCTCTGCGAGCATGACCACTTGTGGCTTAC 1560  
 Db |||||  
 QY 1516 GACCGGCGCACAGCCACACAGCGCTGATCTCTGCGAGCATGACCACTTGTGGCTTAC 1575  
 Db |||||  
 QY 1561 GGCACCTTGCACCCACAGCCACTGAGGACTCGCCAAAGCGCTTCGATGTGAGGCTGCG 1620  
 Db |||||  
 QY 1576 GGCACCTTGCACCCACAGCCACTGAGGACTCGCCAAAGCGCTTCGATGTGAGGCTGCG 1635  
 Db |||||  
 QY 1621 GTGCTGGGCTTCTGAGCGCTTCAAGTGTGCGCTTCACTCTGAGGAGTGTGCTGGCGGAG 1680  
 Db |||||  
 QY 1636 GTGCTGGGCTTCTGAGCGCTTCAAGTGTGCGCTTCACTCTGAGGAGTGTGCTGGCGGAG 1695  
 Db |||||  
 QY 1681 AAGACCCAGTGGGTGATCGGGCTGCGCACAGAAAGCGCAAGCGCGCAAGGCGCATCCAG 1740  
 Db |||||  
 QY 1696 AAGACCCAGTGGGTGATCGGGCTGCGCACAGAAAGCGCAAGCGCGCAAGGCGCATCCAG 1755  
 Db |||||  
 QY 1741 ATCCAGCCAGCGCGGCTTCTATGCAATGCTGATGCAATGCGCAACAGTACAGCGCC 1800  
 Db |||||  
 QY 1756 ATCCAGCCAGCGCGGCTTCTATGCAATGCTGATGCAATGCGCAACAGTACAGCGCC 1815  
 Db |||||  
 QY 1801 TGCAAGGAGCGCTGCAAGCGGCTTAAAGTCCGGGCAAGCTTCAAGTGGGTCTTTC 1860  
 Db |||||  
 QY 1816 TGCAAGGAGCGCTGCAAGCGGCTTAAAGTCCGGGCAAGCTTCAAGTGGGTCTTTC 1875  
 Db |||||  
 QY 1861 CTGGACTATGACCAAGGCTTGTCTCATCTTCTACAATGCTGATGACATGCTGGCTCTAC 1920  
 Db |||||  
 QY 1876 CTGGACTATGACCAAGGCTTGTCTCATCTTCTACAATGCTGATGACATGCTGGCTCTAC 1935  
 Db |||||  
 QY 1921 ACCTTCCGAGAGAGTTCCTGGCAAGCTTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1980  
 Db |||||  
 QY 1936 ACCTTCCGAGAGAGTTCCTGGCAAGCTTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1995  
 Db |||||  
 QY 1981 GCCAATGCAAGAGTTCAGCGCTGCGGATCAACCGTCCGATCTAGTCCAGGGAG 2040  
 Db |||||  
 QY 1996 GCCAATGCAAGAGTTCAGCGCTGCGGATCAACCGTCCGATCTAGTCCAGGGAG 2055  
 Db |||||  
 QY 2041 AAGAGACCAAACTCTCTGGGACCACTGCGCACTGCGAAGCGCTGCCAGAGAGTAG 2100  
 Db |||||  
 QY 2056 AAGAGACCAAACTCTCTGGGACCACTGCGCACTGCGAAGCGCTGCCAGG-AGATAG 2114  
 Db |||||  
 QY 2101 AAGACTCGACTTCCAGCCCACTGGGCACTTGGAGACTTCCAGGCGAGTGTTCCTTCCCTCC 2160  
 Db |||||  
 QY 2115 AAGACTCGACTTCCAGCCCACTGGGCACTTGGAGACTTCCAGGCGAGTGTTCCTTCCCTCC 2174  
 Db |||||  
 QY 2161 AGCTTCCAGTCTGTAAATGAGGTTGATTCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2220  
 Db |||||  
 QY 2175 AGCTTCCAGTCTGTAAATGAGGTTGATTCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2234  
 Db |||||  
 QY 2221 GATGTTCTGTAGCTCTGACCTTGATAGGAGTACAGCTTTGATCCAGGATGTGACATGGC 2280



Query Match Best Local Similarity Matches 119; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	3.1%; Score 119; DB 10; Length 3216;	
	100.0%; Pred. No. 5.3e-54;	
QY	592	GGCCCTGGCGCATGGCTGAGCTCAAGGACGAGCTGCTGTGCTCCATCTGCTGAGC 651
Db	19	GCCCTGGCGCATGGCTGAGCTCAAGGACGAGCTGCTGTGCTCCATCTGCTGAGC 78
QY	652	ATCTACGAGGACCGGCTGAGCTGGCTGGCGAGCACTACTTCTGGCGGCTGCAATCAC 710
Db	79	ATCTACGAGGACCGGCTGAGCTGGCTGGCGAGCACTACTTCTGGCGGCTGCAATCAC 137
RESULT 12		
AK125034/C		
LOCUS	AK125034	3444 bp mRNA linear PRI 09-SEP-2003
DEFINITION	Homo sapiens cDNA FLJ43044 fis, clone BRTHA3003474.	
ACCESSION	AK125034	
VERSION	AK125034.1 GI:34530992	
KEYWORDS	oligo capping; fis (full insert sequence).	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.	
AUTHORS	1 Ota, T., Nakagawa, S., Senoh, A., Mizuguchi, H., Inagaki, H., Sugiyama, T., Irie, R., Orsuki, T., Sato, H., Wakanatsu, A., Ishii, S., Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M., KiKuchi, H., Kanda, K., Wagatsuna, M., Murakawa, K., Kanehori, K., Takahashi-Fujii, A., Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S., Nagahari, K., Masuho, Y., Nagai, K. and Isogai, T.	
TITLE	NEDO human cDNA sequencing project	
JOURNAL	Unpublished	
REFERENCE	2 (bases 1 to 3444)	
AUTHORS	Isogai, T. and Yamamoto, J.	
TITLE	Direct Submission	
JOURNAL	Submitted (15-JUL-2003) Takao Isogai, FLJ Project (HRI Team); 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0818, Japan	
COMMENT	(E-mail: genomics@ri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986) NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan key Technology Center etc.); 5' & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.	
FEATURES	Location/Qualifiers	
source	1..3444	
	/organism="Homo sapiens"	
	/mol_type="mRNA"	
	/db_xref="taxon:9606"	
	/clone="BRTHA3003474"	
	/tissue_type="Chalamus"	
	/clone_lib="BRTHA3"	
	/note="cloning vector: pME18SFL3"	
	6..410	
	/note="unnamed protein product"	
	/codon_start=1	
	/protein_id="BAC86027.1"	
	/db_xref="GI:34530993"	
	/translation="MHWALTGLDGKLPSCHFVAGLVGVPSTSLGSLCTKIEITTVPTWDDFFSLFWFLFLFFIFRWSLTLLVAQAGVQWLDIGLSLQPPPPGFKQFSLSLSSWDYRHPHPPCPAFSFCIFSDVGSFPYFPGNS"	
CDS		
ORIGIN	2.6%; Score 99; DB 9; Length 3444;	
Query Match	100.0%; Pred. No. 8e-43;	
Best Local Similarity	100.0%; Pred. No. 8e-43;	
Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	1009	CAGAGGAGCTGAAGGACCACTTCAGGCCCTTCAAGACAGCGAGCGGGAACACCCGAA 1068
Db	1	
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov		
Series: Plate: Row: Column: 0		
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis.		
Location/Qualifiers		
1..4150		
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov		
Series: Plate: Row: Column: 0		
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis.		
Location/Qualifiers		
1..4150		

```

/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:6831932"
/tissue_type="Brain, mouse, 13.5,14.5,16.5,17.5 dpc"
/clone_lib="NIH_BMAP_FY0"
/lab_host="DH10B"
/note="Vector: pYX-ASC"

ORIGIN
Query Match          0.7%; Score 27; DB 10; Length 4150;
Best Local Similarity 100.0%; Pred. No. 0.014;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3800 AGATGTCAAAAAAGAAAAAAGAAAAA 3826
Db 4111 AGATGTCAAAAAAGAAAAAAGAAAAA 4137

RESULT 14
HSM803417
LOCUS HSM803417 4685 bp mRNA linear PRI 13-MAY-2003
DEFINITION Homo sapiens mRNA; cDNA DKFZp313H168 (from clone DKFZp313H168).
ACCESSION AL832110
VERSION AL832110.1 GI:21732653
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 4685)
Wambutt,R., Heubner,D., Mewes,H.W., Weil,B., Amid,C., Osanger,A.,
Robo,G., Han,M. and Wiemann,S.
Direct Submission
Submitted (13-MAY-2003) MIPS, Ingolstaedter Landstr.1, D-85764
Neuherberg, GERMANY
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by AGOWA (Berlin/Germany) within the cDNA sequencing
consortium of the German Genome Project.
This clone (DKFZp313H168) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
information about the clone and the sequencing project is available
at http://mips.gsf.de/proj/cDNA/.

FEATURES
source
1..4685
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp313H168"
/tissue_type="human bone marrow"
/clone_lib="313 (synonym: hlcc2). Vector pTriplex2; host
DH10B; sites SfilA + SfilB"
/dev_stage="adult"
polyA_signal 4647..4652
polyA_site 4662

ORIGIN
Query Match          0.7%; Score 27; DB 9; Length 4685;
Best Local Similarity 100.0%; Pred. No. 0.014;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3800 AGATGTCAAAAAAGAAAAAAGAAAAA 3826
Db 4656 AGATGTCAAAAAAGAAAAAAGAAAAA 4682

RESULT 15
BC019375
LOCUS BC019375 2619 bp mRNA linear ROD 04-OCT-2003
DEFINITION Mus musculus DNA segment, Chr 1, ERATO Doi 622, expressed, mRNA

```

```

(cDNA clone MGC:30533 IMAGE:5008599), complete cds.
BC019375
VERSION BC019375.1 GI:18043921
KEYWORDS MGC.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 2619)
Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Ustin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Adams,R.D., Mullany,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Morley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodriguez,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smallos,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
22388257
12477932
2 (bases 1 to 2619)
Strausberg,R.
Direct Submission
Submitted (19-DEC-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgaps-rc@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amg@bcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H.,
Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
A.N., Gibbs, R.A.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 41 Row: k Column: 6
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 19527041.

FEATURES
source
1..2619
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N-3"
/db_xref="taxon:10090"
/clone="MGC:30533 IMAGE:5008599"
/tissue_type="Mammary tumor. MMTV-LTR/INT3 model. 5 month
old mouse. Taken by biopsy."
/clone_lib="NCI CGAP_Mam2"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
1..2619
/gene="D1Ertd622e"

```

```
/db_xref="LocusID:52392"
/db_xref="MGI:1277184"
159...782
/codon_start=1
/product="DlRtd622e protein"
/protein_id="AAH19375.1"
/db_xref="GI:18043922"
/db_xref="LocusID:52392"
/translation="MEVDINGDSRSTLTPLVAEGSSPGKAEKPCSTPCSPM
RTVSGYQLHMDSNYLVGFTIGBELLKLAQKCTGGEDSKGEAMPALRAKQIDIGLAR
SSRLYKTSRYQPYEIPAVNGRRRRMPSSGDKCTKPLPYEPYKALHGPPLCLLKG
KRAHSKSLDYLNLDKMNIKEPADTEVLQYQLQLTLRGDRVFARNNT"
```

ORIGIN

```
Query Match      0.7%; Score 26; DB 10; Length 2619;
Best Local Similarity 100.0%; Pred.No. 0.051;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3801 GATGTCAAAAAAAAAAAAAAAAAAAA 3826
      |||||||
Db 2583 GATGTCAAAAAAAAAAAAAAAAAAAA 2608
```

Search completed: September 4, 2004, 11:29:30  
Job time : 9627 secs



This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 4, 2004, 06:55:12 ; Search time 944 Seconds  
(without alignments)  
17217.803 Million cell updates/sec

Title: US-09-927-091-3

Perfect score: 3826

Sequence: 1 aggtcgctggaccgaagc.....aaaaaaaaaaaaaaaaaaaaa 3826

Scoring table: OLIGO NUC

Gapop\_60.0 , Gapext 60.0

Searched: 3373863 seqs, 2124099041 residues

Word size : 18

Total number of hits satisfying chosen parameters: 22942

Minimum DB seq length: 2500

Maximum DB seq length: 5000

Post-processing: Listing first 45 summaries

Database : N\_Geneseqn\_29Jan04.\*  
1: Geneseqn1980s.\*  
2: Geneseqn1990s.\*  
3: Geneseqn2000s.\*  
4: Geneseqn2001as.\*  
5: Geneseqn2001bs.\*  
6: Geneseqn2001cs.\*  
7: Geneseqn2002as.\*  
8: Geneseqn2003bs.\*  
9: Geneseqn2003cs.\*  
10: Geneseqn2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3826	100.0	3826	6	Abk12806 Human cDN
2	3163	82.7	3436	4	Aah14509 Human cDN
3	2994	78.3	3243	9	Adc37268 Nuclear f
4	26	0.7	4702	7	Abx76356 Lung canc
5	25	0.7	2522	2	Aav82781 Clone df3
6	25	0.7	2522	6	Abq92018 Human pol
7	25	0.7	2586	9	Adb62721 Human cDN
8	25	0.7	2629	5	Aah34824 Human col
9	25	0.7	2647	5	Abv23296 Human pro
10	25	0.7	2647	5	Abv25449 Human pro
11	25	0.7	2647	5	Abv30164 Human pro
12	25	0.7	2647	5	Abv25571 Human pro
13	25	0.7	2647	5	Abv29143 Human pro
14	25	0.7	2796	2	Aax35727 cDNA enco
15	25	0.7	3858	9	Adb69641 Human REM
16	25	0.7	3891	3	Aac76424 Human ORF
17	25	0.7	4493	8	Ac98971 cDNA enco
18	25	0.7	4498	8	Ada38053 DNA relat
19	25	0.7	4572	6	Abq98201 Human ost
20	25	0.7	4595	4	Ab115850 Human ost
21	25	0.7	4717	6	Abk99971 DNA enco
22	25	0.7	4729	2	Aaq04043 Recombina
23	25	0.7	4740	3	Aaa35032 Human ade

24	25	0.7	4740	3	AAF21154 Human low
25	25	0.7	4740	6	ABK50046 cDNA enco
26	25	0.7	4740	6	ABK83766 Human cDN
27	25	0.7	4740	7	ABZ96848 Human nuc
28	25	0.7	4740	7	ABV76142 Human Mac
29	25	0.7	4740	9	ADD25614 Binding d
30	24	0.6	2509	3	Aac78146 Human can
31	24	0.6	2512	9	Adc87242 Human pan
32	24	0.6	2558	4	ABL21586 Drosophil
33	24	0.6	2612	4	AAD05059 Human sec
34	24	0.6	2612	7	ADA56044 Gene enco
35	24	0.6	2612	7	ADA39854 Human sec
36	24	0.6	2633	4	AAL35789 Human mus
37	24	0.6	2633	7	ABX58777 cDNA enco
38	24	0.6	2642	4	ABL10304 Drosophil
39	24	0.6	2645	7	ABX63799 Human cDN
40	24	0.6	2661	7	ABZ18582 Group III
41	24	0.6	2677	4	AAD18016 Coniothyr
42	24	0.6	2690	6	ABA93708 Human tra
43	24	0.6	2720	4	AAH18311 Human cDN
44	24	0.6	2723	4	ABL27612 Drosophil
45	24	0.6	2820	3	AAA23433 cDNA enco

ALIGNMENTS

RESULT 1  
ABK12806  
ID ABK12806 standard; cDNA; 3826 BP.  
XX  
AC ABK12806;  
XX  
DT 18-JUN-2002 (first entry)  
XX  
DE Human cDNA encoding tumour suppressor CAR-1.  
XX

Human; ss; tumour suppressor; gene; CAR-1; cytostatic; cancer; tumour; gene therapy; brain cancer; lung cancer; liver cancer; kidney cancer; colon cancer; stomach cancer; breast cancer; endometrial cancer; prostate cancer; testicular cancer; ovarian cancer; skin cancer; head and neck cancer; oesophageal cancer; bone marrow cancer; chromosome 1p31-1p36.  
XX  
OS Homo sapiens.  
XX

Key Location/Qualifiers  
CDS 604..2031  
FT /\*tag= a  
FT /product= "CAR-1"  
FT  
WO200212285-A2.  
14-FEB-2002.  
09-AUG-2001; 2001WO-US025269.  
10-AUG-2000; 2000US-0225033P.  
23-AUG-2000; 2000US-0227560P.  
(TEXA ) UNIV TEXAS SYSTEM.  
Killary A, Chandler D, Lott S;  
WPI; 2002-269088/31.  
P-PSDB; AAU78657.  
New tumor suppressor CAR-1 polypeptides and polynucleotides, useful for diagnosing cancer, for altering the phenotype of a tumor cell, for treating cancers or as a diagnostic or prognostic indicator of cancer.  
Claim 2; Page 134-135; 185pp; English.  
XX



Db 1501 GACCGGGCACAGCCACACAGCGCCTGATCTGTGCGAGACTGCACCACTTGTGGCTTAC 1560  
QY 1561 GSCAACTTGCACCACAGACCACTGCAGGACTCGCBAAGCGTTTCGATGTGAGGTGTGCG 1620  
Db 1561 GSCAACTTGCACCACAGACCACTGCAGGACTCGCBAAGCGTTTCGATGTGAGGTGTGCG 1620  
QY 1621 GTGCTGGGTTCCTGAAGCCTTCAGTAGTGGCTGCCACTACTGTGGAGGTGGTGGCGGAG 1680  
Db 1621 GTGCTGGGTTCCTGAAGCCTTCAGTAGTGGCTGCCACTACTGTGGAGGTGGTGGCGGAG 1680  
QY 1681 AAGACCCAGTGGGTGATCGGGCTGGCACACGAAAGCGCGAAGCGCGAAGCGAGCATCCAG 1740  
Db 1681 AAGACCCAGTGGGTGATCGGGCTGGCACACGAAAGCGCGAAGCGCGAAGCGAGCATCCAG 1740  
QY 1741 ATCCAGCCAGCGCGGCTTCTACTGCATCGTGATGCAAGTGGCAACCAAGTACAGCGCC 1800  
Db 1741 ATCCAGCCAGCGCGGCTTCTACTGCATCGTGATGCAAGTGGCAACCAAGTACAGCGCC 1800  
QY 1801 TGCACGGAGCCCTGGACGGGCTTAACTGCGGGACAAGCTTGACAAGGTGGGTGTCTTC 1860  
Db 1801 TGCACGGAGCCCTGGACGGGCTTAACTGCGGGACAAGCTTGACAAGGTGGGTGTCTTC 1860  
QY 1861 CTGGAATATGACAAGGCTTGTCTCATCTCAATGCTGATGACATGTCTGGCTCTAC 1920  
Db 1861 CTGGAATATGACAAGGCTTGTCTCATCTCAATGCTGATGACATGTCTGGCTCTAC 1920  
QY 1921 ACCTTCCGGAGAGTTCCCTGGAGCTCTGCTCTTACTTCAAGCCTGAGCCCTGGCAGGCCAC 1980  
Db 1921 ACCTTCCGGAGAGTTCCCTGGAGCTCTGCTCTTACTTCAAGCCTGAGCCCTGGCAGGCCAC 1980  
QY 1981 GSCAATGGCAAGACTTTCAGCGCTGCGGATCAACACCGTCCGATCTAGTCCAGGCAG 2040  
Db 1981 GSCAATGGCAAGACTTTCAGCGCTGCGGATCAACACCGTCCGATCTAGTCCAGGCAG 2040  
QY 2041 AAGGAGACCAACCTCTCTGGACCACTGCCACTGCAAGAGCCCTGCCAGGAAGATAG 2100  
Db 2041 AAGGAGACCAACCTCTCTGGACCACTGCCACTGCAAGAGCCCTGCCAGGAAGATAG 2100  
QY 2101 AAGACTGACCTCCAGCCGACCGTGGCCACTGGAGACTCAGGCCAGTGTGTACCTCC 2160  
Db 2101 AAGACTGACCTCCAGCCGACCGTGGCCACTGGAGACTCAGGCCAGTGTGTACCTCC 2160  
QY 2161 AGCCTCCAGTCTGTAATAAGAGGTGCACTCCCTACTTCCCTAAACTCTCTTCCAGCATC 2220  
Db 2161 AGCCTCCAGTCTGTAATAAGAGGTGCACTCCCTACTTCCCTAAACTCTCTTCCAGCATC 2220  
QY 2221 GATGTTCTGATGCTCTGACCTTGATAGGATACAGCTTTGATCCAAAGATGTGACATGGC 2280  
Db 2221 GATGTTCTGATGCTCTGACCTTGATAGGATACAGCTTTGATCCAAAGATGTGACATGGC 2280  
QY 2281 TTCTCTCAGGGCAACCCCTGCCACCTCATCCCATCTTCTCAGGGCGAGGGACTA 2340  
Db 2281 TTCTCTCAGGGCAACCCCTGCCACCTCATCCCATCTTCTCAGGGCGAGGGACTA 2340  
QY 2341 CCTTCCAGTGTCTCCCTCCAGCCAGCCCTGACCTCAGGAAGTGTCCAGAGCATGGCCAGT 2400  
Db 2341 CCTTCCAGTGTCTCCCTCCAGCCAGCCCTGACCTCAGGAAGTGTCCAGAGCATGGCCAGT 2400  
QY 2401 AGTTGCGACCGGAAAGACACAGACCCCTCTTATGTCCCATGGCTAAGACTTACCCC 2460  
Db 2401 AGTTGCGACCGGAAAGACACAGACCCCTCTTATGTCCCATGGCTAAGACTTACCCC 2460  
QY 2461 TGACCAAGCTAGTAGTGGGCCATTTACCTTGACCCCTGACAGTGGTCAAGTAGT 2520  
Db 2461 TGACCAAGCTAGTAGTGGGCCATTTACCTTGACCCCTGACAGTGGTCAAGTAGT 2520  
QY 2521 ACCTGTCTTAGGGTGTGAGAGCAACCTCTCTCGCCACCCGCCACCAAGAACTAT 2580  
Db 2521 ACCTGTCTTAGGGTGTGAGAGCAACCTCTCTCGCCACCCGCCACCAAGAACTAT 2580  
QY 2581 ATGGTTCCTACTTCTCCCACTGATCTGCTGGTCAAGTATGCTGTGGCCTGTGGAAGG 2640  
Db 2581 ATGGTTCCTACTTCTCCCACTGATCTGCTGGTCAAGTATGCTGTGGCCTGTGGAAGG 2640

QY 2641 CACCTGTTAGTTGAGTCCACACATTATAGTGTGCGACCACTTCTGCGCCACAGGCC 2700  
Db 2641 CACCTGTTAGTTGAGTCCACACATTATAGTGTGCGACCACTTCTGCGCCACAGGCC 2700  
QY 2701 GAGGACAGGGTGTAGGGTATACCCAAAGCTGATGACAGAGCCCAATTAGCCTAAAGCAACT 2760  
Db 2701 GAGGACAGGGTGTAGGGTATACCCAAAGCTGATGACAGAGCCCAATTAGCCTAAAGCAACT 2760  
QY 2761 GAGGACAGAGCTTCCCTGATGATGAGGTCCCGAGTAGCTCTGAACAAGTCCAGGCCA 2820  
Db 2761 GAGGACAGAGCTTCCCTGATGATGAGGTCCCGAGTAGCTCTGAACAAGTCCAGGCCA 2820  
QY 2821 ACCCTCTTCAGCGACGCCCTCTGTGACCTGAGGTGACAGGAGCTTCCAGAGCAGTTG 2880  
Db 2821 ACCCTCTTCAGCGAGGCCCTCTGTGACCTGAGGTGACAGGAGCTTCCAGAGCAGTTG 2880  
QY 2881 TTGTAATTAGGACCCCAAGCACTGGGAGGGGCTGTTGGCTAGACCCCTTGTCAACTTGGC 2940  
Db 2881 TTGTAATTAGGACCCCAAGCACTGGGAGGGGCTGTTGGCTAGACCCCTTGTCAACTTGGC 2940  
QY 2941 ATCTATCTCAGTTAGGATCTGCTGCAGAAAAACAAGAGCCACTTGTAGCTGTGTTAAATTA 3000  
Db 2941 ATCTATCTCAGTTAGGATCTGCTGCAGAAAAACAAGAGCCACTTGTAGCTGTGTTAAATTA 3000  
QY 3001 GACAAAGATTACTACCTGGCCCCCTGGTGGCTTGCAAAATTTGTTGAAAGAGCTGGAGAAG 3060  
Db 3001 GACAAAGATTACTACCTGGCCCCCTGGTGGCTTGCAAAATTTGTTGAAAGAGCTGGAGAAG 3060  
QY 3061 CAGACTCTGCTGAATTTCCAGGAACTCCAGCGGCCAGATTTCATCATGTCTGTGTGACCA 3120  
Db 3061 CAGACTCTGCTGAATTTCCAGGAACTCCAGCGGCCAGATTTCATCATGTCTGTGTGACCA 3120  
QY 3121 GGAAGCTGCCCCCATCTGCAGGAAGCCACTATATGCGAAGAAAGTGTGACTGCGAGAACTA 3180  
Db 3121 GGAAGCTGCCCCCATCTGCAGGAAGCCACTATATGCGAAGAAAGTGTGACTGCGAGAACTA 3180  
QY 3181 GGTCTCCTCTGCGACGGTCCGTGCAGCCAAATAGATGTCTGAGGCTCTGCCCTCTCCCA 3240  
Db 3181 GGTCTCCTCTGCGACGGTCCGTGCAGCCAAATAGATGTCTGAGGCTCTGCCCTCTCCCA 3240  
QY 3241 CTTCACTCAGTTCCCAAAATCTAAATTTTACAGAGATTCTGTTGGGGGAACTTAAGTC 3300  
Db 3241 CTTCACTCAGTTCCCAAAATCTAAATTTTACAGAGATTCTGTTGGGGGAACTTAAGTC 3300  
QY 3301 AGATCCAGAACTTGGCTGCAAGGAGTCTGGAATGTCAATTTCCCTAGAGGAAGTTA 3360  
Db 3301 AGATCCAGAACTTGGCTGCAAGGAGTCTGGAATGTCAATTTCCCTAGAGGAAGTTA 3360  
QY 3361 GGTGGGTGGAGCAAGCCCACTGCGTTTTCGCCACAGCATCCAATCGTGAAGAACT 3420  
Db 3361 GGTGGGTGGAGCAAGCCCACTGCGTTTTCGCCACAGCATCCAATCGTGAAGAACT 3420  
QY 3421 CGGAGAGAGGTGGAGTCCACATCTAGGGTGTCTGCCCCCTTGGCTCTATCCCTGCCCCAG 3480  
Db 3421 CGGAGAGAGGTGGAGTCCACATCTAGGGTGTCTGCCCCCTTGGCTCTATCCCTGCCCCAG 3480  
QY 3481 AGGTGGAACTGGAGAGTGGGCTGGAAGCTGAGCCTTAAATGTCTCCCGGCTTGACT 3540  
Db 3481 AGGTGGAACTGGAGAGTGGGCTGGAAGCTGAGCCTTAAATGTCTCCCGGCTTGACT 3540  
QY 3541 TTTCTTTCTAGTCTTGGGGCTAGATTCTGCACTTTGGGTCTCTGACACAACACACCATC 3600  
Db 3541 TTTCTTTCTAGTCTTGGGGCTAGATTCTGCACTTTGGGTCTCTGACACAACACACCATC 3600  
QY 3601 CCAAGTAGTCCGGAAGAGCTAAACAGGGGTCTTAAATGGCTGCCCGCCACCCG 3660  
Db 3601 CCAAGTAGTCCGGAAGAGCTAAACAGGGGTCTTAAATGGCTGCCCGCCACCCG 3660  
QY 3661 GGCCTCCCTTGGGCAAAAGGAATTTGTAGCCCTTACCCCAACCTTCACTACAGAACT 3720  
Db 3661 GGCCTCCCTTGGGCAAAAGGAATTTGTAGCCCTTACCCCAACCTTCACTACAGAACT 3720

QY 3721 GGGCCACCCAGAGTATTTTATTTAAATGTTGCCATTTTATGAGTTATGATCAATT 3780  
 Db |||||  
 QY 3721 GGGCCACCCAGAGTATTTTATTTAAATGTTGCCATTTTATGAGTTATGATCAATT 3780  
 Db |||||  
 QY 3781 TGTATTAAATTAAGTTTACAGATGCAAAAAAATAAAAAA 3826  
 Db 3781 TGTATTAAATTAAGTTTACAGATGCAAAAAAATAAAAAA 3826

RESULT 2  
 AAH14509  
 ID AAH14509 standard; cDNA; 3436 BP.  
 XX  
 AC AAH14509;  
 XX  
 XX 26-JUN-2001 (first entry)  
 XX  
 DE Human cDNA sequence SEQ ID NO:12036.  
 XX  
 XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN EP1074617-A2.  
 XX  
 PD 07-FEB-2001.  
 XX  
 XX 28-JUN-2000; 2000EP-00116126.  
 XX  
 PR 29-JUL-1999; 99JP-00248036.  
 PR 27-AUG-1999; 99JP-00300253.  
 PR 11-JAN-2000; 2000JP-00118776.  
 PR 02-MAY-2000; 2000JP-00183767.  
 PR 09-JUN-2000; 2000JP-00241899.  
 XX  
 XX (HELI-) HELIX RES INST.  
 XX  
 XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
 XX  
 DR WPI; 2001-318749/34.  
 XX  
 XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.

PS Claim 8; SEQ ID NO 12036; 2537pp + Sequence Listing; English.  
 XX  
 CC The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises:  
 CC (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the  
 CC of an oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the  
 CC oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesizing polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialized methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention  
 CC  
 XX Sequence 3436 BP; 731 A; 1126 C; 901 G; 678 T; 0 U; 0 Other;

Query Match 82.7%; Score 3163; DB 4; Length 3436;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 3433; Conservative 0; Mismatches 3; Indels 1; Gaps 1;  
 QY 370 CCACCTACTCTAGACTGCCCTCCGGGCTGGCGTCCACCGAGTCTCAGCGCGCACCCC 429  
 Db 1 CCACCTACTCTAGACTGCCCTCCGGGCTGGCGTCCACCGAGTCTCAGCGCGCACCCC 60  
 QY 430 TTCTCGCGTTACCTCTCTCGGAGACAGACCCCTCTCTCTCGGTAGTCTCTACCCC 489  
 Db 61 TTCTCGCGTTACCTCTCTCGGAGACAGACCCCTCTCTCTCGGTAGTCTCTACCCC 120  
 QY 490 TGCCTGTGGGGCTCTGTCCTCCCGCGCCAGCCCTCGGTGCTCTCCGACAGCGCGCG 549  
 Db 121 TGCCTGTGGGGCTCTGTCCTCCCGCGCCAGCCCTCGGTGCTCTCCGACAGCGCGCG 180  
 QY 550 TCTCTCAGCCGCCCTCTGCCCCCTCGGGGCCCTCTCTCTGCTGCCCCCTGGCGCATGGCG 609  
 Db 181 TCTCTCAGCCGCCCTCTGCCCCCTCGGGGCCCTCTCTCTGCTGCCCCCTGGCGCATGGCG 240  
 QY 610 TGCAGCCTCAAGGACGAGCTGCTGCTGCTCCATCTGCGTGCATCTACCGAGACCCCGTG 669  
 Db 241 TGCAGCCTCAAGGACGAGCTGCTGCTGCTCCATCTGCGTGCATCTACCGAGACCCCGTG 300  
 QY 670 AGCCTGGGCTGCGAGCAGCTACTTCTGCGCGCGCTGCATCAGGAGCACTGGGTGGCGAG 729  
 Db 301 AGCCTGGGCTGCGAGCAGCTACTTCTGCGCGCGCTGCATCAGGAGCACTGGGTGGCGAG 360  
 QY 730 GAGCGCAGCGCGCCCGCGACTGCCCCAGTGTGCGCGCGCACTTGTGCGAGCGCGCGCTG 789  
 Db 361 GAGCGCAGCGCGCCCGCGACTGCCCCAGTGTGCGCGCGCACTTGTGCGAGCGCGCGCTG 420  
 QY 790 GCGCCGAGCCTCAAGCTGGCCCAATCGTGGAGCGCTACAGTCTCTTCCCGCTGAGCGCC 849  
 Db 421 GCGCCGAGCCTCAAGCTGGCCCAATCGTGGAGCGCTACAGTCTCTTCCCGCTGAGCGCC 480  
 QY 850 ATCTCTAAACCGCGCGCGCGCGCCCTGCGCGCGCACCGACGACAAAGGTCAAGTCTTTC 909  
 Db 481 ATCTCTAAACCGCGCGCGCGCGCCCTGCGCGCGCACCGACGACAAAGGTCAAGTCTTTC 540  
 QY 910 TGCCTCAGGACCGCGCGCTTCTGCTTTCTTGTGCGAGCGCTGCACTGCACGAGCAG 969  
 Db 541 TGCCTCAGGACCGCGCGCTTCTGCTTTCTTGTGCGAGCGCTGCACTGCACGAGCAG 600  
 QY 970 CATCAGTTCACCGCATCGAGCGCTTGCAGCGCTGCGAGGAGGAGCTGAAGSCCAA 1029  
 Db 601 CATCAGTTCACCGCATCGAGCGCTTGCAGCGCTGCGAGGAGGAGCTGAAGSCCAA 660  
 QY 1030 CTTTCAGGCGCTTCAAGACAGCGAGCGGGAACACACCGAAGCGCTGCAGTCTCTCAAGCGA 1089  
 Db 661 CTTTCAGGCGCTTCAAGACAGCGAGCGGGAACACACCGAAGCGCTGCAGTCTCTCAAGCGA 720  
 QY 1090 CAATGCGGAGACCAAGTCTTCCACCAAGAGCGCTGCGGACCACTATGCGGAGGCGCTTC 1149  
 Db 721 CAATGCGGAGACCAAGTCTTCCACCAAGAGCGCTGCGGACCACTATGCGGAGGCGCTTC 780  
 QY 1150 GAGCGCTGCACCGCGCTGCTGCGTGAACCGCAGAGGCGCATGTAGAGGAGCTGGAGCG 1209  
 Db 781 GAGCGCTGCACCGCGCTGCTGCGTGAACCGCAGAGGCGCATGTAGAGGAGCTGGAGCG 840  
 QY 1210 GACACGCGCGCGCTGAGCGCATCGAGCAGAAAGTCCAGCGCTTACGCCAGCAGCTG 1269  
 Db 841 GACACGCGCGCGCTGAGCGCATCGAGCAGAAAGTCCAGCGCTTACGCCAGCAGCTG 900  
 QY 1270 CGAAGTTCAGGAGGAGCGCCAGATCTCTGAGAGCGGCTGGCTGAACCGACCGGCGAC 1329  
 Db 901 CGAAGTTCAGGAGGAGCGCCAGATCTCTGAGAGCGGCTGGCTGAACCGACCGGCGAC 960  
 QY 1330 ACCTTCTGCTGGGTGGGCTCTCACTGTCCGAGCGGCTCAAGGGGAAAAATCCATGAGACC 1389  
 Db 961 ACCTTCTGCTGGGTGGGCTCTCACTGTCCGAGCGGCTCAAGGGGAAAAATCCATGAGACC 1020

QY	1390	AACCTCAATATGAAGACTTCCCGACCTCCAAGTACACAGGCCCTTCGAGTACACCATC	1440
Db	1021	AACCTCACATATAAGACTTCCCGACCTCCAAGTACACAGGCCCTTCGAGTACACCATC	1080
QY	1450	TGGAAGTCCCTGTTCAGGACATCCACCAGTGCCAGCGCCTTAACCTTGGACCCGGGC	1509
Db	1081	TGGAAGTCCCTGTTCAGGACATCCACCAGTGCCAGCGCCTTAACCTTGGACCCGGGC	1140
QY	1510	ACAGCCACACAGGCGCTGATCTGTTCGAGCGACTGCAACCATTTGGCTTACGGCAACTTG	1569
Db	1141	ACAGCCACACAGGCGCTGATCTGTTCGAGCGACTGCAACCATTTGGCTTACGGCAACTTG	1200
QY	1570	CACCCACAGCACTGCAGGACTCGCCAAAGCGTTTCGATGTGAGGTGTTCGGTCTGGGT	1629
Db	1201	CACCCACAGCACTGCAGGACTCGCCAAAGCGTTTCGATGTGAGGTGTTCGGTCTGGGT	1260
QY	1630	TCTGAGCCTTCAGTAGTGCGCTCCACTACTGGGAGGTGTGTGGCGGAGAAGACCCAG	1689
Db	1261	TCTGAGCCTTCAGTAGTGCGCTCCACTACTGGGAGGTGTGTGGCGGAGAAGACCCAG	1320
QY	1690	TGGGTGATCGGGGTGGCACGAAGCGCGAAGCGCAAGCGCAAGCGCAAGCATCCAGCCCC	1749
Db	1321	TGGGTGATCGGGGTGGCACGAAGCGCGAAGCGCGAAGCGCAAGCATCCAGCCCC	1380
QY	1750	AGCGCGGCTTCTACTGCTGATGTCAGATGCGAACCGAGTACAGCGCTGCACGGAG	1809
Db	1381	AGCGCGGCTTCTACTGCTGATGTCAGATGCGAACCGAGTACAGCGCTGCACGGAG	1440
QY	1810	CCCTGGACGGCTTAACGTTCGGGACAAGCTTTGACAGGTGGGTGTCTTCTGGACTAT	1869
Db	1441	CCCTGGACGGCTTAACGTTCGGGACAAGCTTTGACAGGTGGGTGTCTTCTGGACTAT	1500
QY	1870	GACCAAGCTTGCTCACTTCTCAATGCTGATGACATGTCTGGCTCTACACTTCCGC	1929
Db	1501	GACCAAGCTTGCTCACTTCTCAATGCTGATGACATGTCTGGCTCTACACTTCCGC	1560
QY	1930	GAGAAGTTCCTGGCAAGCTTGCTCTTAATTACAGCCTGCCAGAGCCACGCCAATGTC	1989
Db	1561	GAGAAGTTCCTGGCAAGCTTGCTCTTAATTACAGCCTGCCAGAGCCACGCCAATGTC	1620
QY	1990	AAGAACTTCAGCGCTGGGATCAACACCGTTCGGATCTAGTCCAGGCAAGAGGAGCC	2049
Db	1621	AAGAACTTCAGCGCTGGGATCAACACCGTTCGGATCTAGTCCAGGCAAGAGGAGCC	1680
QY	2050	ACAACTCTCTGGGACACTGCCACTTCAAAGAGCCCTGCCACGAAGATAGAAGACCTGG	2109
Db	1681	ACAACTCTCTGGGACACTGCCACTTCAAAGAGCCCTGCCACGAAGATAGAAGACCTGG	1739
QY	2110	ACTCCAGCCCAACCGTGGCCACTGGAGACCTCAGGCCAGTTGTTTACCTCCAGCTCCAG	2169
Db	1740	ACTCCAGCCCAACCGTGGCCACTGGAGACCTCAGGCCAGTTGTTTACCTCCAGCTCCAG	1799
QY	2170	TCTGTAAAATGGAGGTGCAATCCCTACTTCTTAAACTCTCTCCAGCATCGATGTCTG	2229
Db	1800	TCTGTAAAATGGAGGTGCAATCCCTACTTCTTAAACTCTCTCCAGCATCGATGTCTG	1859
QY	2230	TAGCTCTGACTCTGATAGGATACAGCTTTGATCCAAAGATGTGACATGCTTCTCTCTCA	2289
Db	1860	TAGCTCTGACTCTGATAGGATACAGCTTTGATCCAAAGATGTGACATGCTTCTCTCTCA	1919
QY	2290	GGGCAACCCCTGCCCAACCCCTCATCCCCTCTTCTCAGGGGCGAGGACTACCTTCCAGT	2349
Db	1920	GGGCAACCCCTGCCCAACCCCTCATCCCCTCTTCTCAGGGGCGAGGACTACCTTCCAGT	1979
QY	2350	GTCTCCCTCCAGCCCGAGCTTGAGTTCAGGAAGTTCAGAGCATGGCCAGTAGTTGCGAG	2409
Db	1980	GTCTCCCTCCAGCCCGAGCTTGAGTTCAGGAAGTTCAGAGCATGGCCAGTAGTTGCGAG	2039
QY	2410	CCGGAAGACACACAGCACCTCTTATGTCCCATGGCTTAAGACTTACCCTGACCAAGC	2469
Db	2040	CCGGAAGACACACAGCACCTCTTATGTCCCATGGCTTAAGACTTACCCTGACCAAGC	2099
QY	2470	TAGTGATGGGCCATTATACCTTTGACCCCAAGTCCACAGTGTTCACAGGTAGTACTGTGTC	2529

Db	2100	TAGTGATGGGCCAATTACCTTGTACCCAGTCCACAGTGGTCCAGGTAGTACCTGGTCC	2159
QY	2530	TAGGGTTGGCTTGAGAGCAACCTCTCTGTCGCCACCCACACCAAGAACAATATATGGTTCCCT	2589
Db	2160	TAGGGTTGGCTTGAGAGCAACCTCTCTGTCGCCACCCACACCAAGAACAATATATGGTTCCCT	2219
QY	2590	ACTTCTCCCACTGATCTGCTGGTCAGTGATGCTGTGGCCCTGTGGAAGCACCTGGTA	2649
Db	2220	ACTTCTCCCACTGATCTGCTGGTCAGTGATGCTGTGGCCCTGTGGAAGCACCTGGTA	2279
QY	2650	GTTGAGTCCACACATATATAGTCATGTGCCACACCTTCTGTGCCACACAGGCCGAGGGACAG	2709
Db	2280	GTTGAGTCCACACATATATAGTCATGTGCCACACCTTCTGTGCCACACAGGCCGAGGGACAG	2339
QY	2710	GGTGAGGGTATACCCAAAGCTGATGCAGAGCCCATTAAGCCCTAAAGCACTGCAGAGCAA	2769
Db	2340	GGTGAGGGTATACCCAAAGCTGATGCAGAGCCCATTAAGCCCTAAAGCACTGCAGAGCAA	2399
QY	2770	GCCTCCCTGATGATCGAGGTCCCACTAGTCTGAAACAAGAGTCCAGCCAAACCTCTTC	2829
Db	2400	GCCTCCCTGATGATCGAGGTCCCACTAGTCTGAAACAAGAGTCCAGCCAAACCTCTTC	2459
QY	2830	AGCCAGGCTCTGTGACCTGCTAGGTGCAAGGCTTCCAGAGCAGTGTGTTGTAATTA	2889
Db	2460	AGCCAGGCTCTGTGACCTGCTAGGTGCAAGGCTTCCAGAGCAGTGTGTTGTAATTA	2519
QY	2890	GGACCCAAAGCACTGGGAGGGCTGTGGCTAGACCCCTGTGCAGCTTGGCATCTATCTC	2949
Db	2520	GGACCCAAAGCACTGGGAGGGCTGTGGCTAGACCCCTGTGCAGCTTGGCATCTATCTC	2579
QY	2950	AGTTAGGATCTGTCTGCAGAAAAACAAGGCCACTTGTAGCTGGTTTAAITAGACAAGGAT	3009
Db	2580	AGTTAGGATCTGTCTGCAGAAAAACAAGGCCACTTGTAGCTGGTTTAAITAGACAAGGAT	2639
QY	3010	TTACTACTGGCCCTGGTGGCTTGCAAAATTTGTTGGAAGAGCTGGAAGAAGCAGTCTG	3069
Db	2640	TTACTACTGGCCCTGGTGGCTTGCAAAATTTGTTGGAAGAGCTGGAAGAAGCAGTCTG	2699
QY	3070	CTGAAATTCAGGAACTCCAGCGCCAGATTCATCATGCTGTTGTGACACAGAAAGCTG	3129
Db	2700	CTGAAATTCAGGAACTCCAGCGCCAGATTCATCATGCTGTTGTGACACAGAAAGCTG	2759
QY	3130	CCCCATCTGCAGGAAGCCACTATGCCAGAAAGCTGCTGACTGCAGAACTAGGCTCCCTC	3189
Db	2760	CCCCATCTGCAGGAAGCCACTATGCCAGAAAGCTGCTGACTGCAGAACTAGGCTCCCTC	2819
QY	3190	TGCCACGGTCCGTGCCAGCAATAGATGCTGTAGGCTCGCCCTCTCCCACTTCACTCA	3249
Db	2820	TGCCACGGTCCGTGCCAGCAATAGATGCTGTAGGCTCGCCCTCTCCCACTTCACTCA	2879
QY	3250	GTTTCCCAAACTAAATTTTACAGAGATTCGTGTTGGGGGAACTTAAGTCAGATCCAGA	3309
Db	2880	GTTTCCCAAACTAAATTTTACAGAGATTCGTGTTGGGGGAACTTAAGTCAGATCCAGA	2939
QY	3310	ACCTTGGCTGCAGGGAGTCTGGGAAATGTCAATTCCTAGAGGAAAGTTAGGGTGGGTG	3369
Db	2940	ACCTTGGCTGCAGGGAGTCTGGGAAATGTCAATTCCTAGAGGAAAGTTAGGGTGGGTG	2999
QY	3370	GAGCAAGCCCACTGCGTTTTCCTGCCACAGCATCCAAATCGTGAAGAACTCGGGAGAGG	3429
Db	3000	GAGCAAGCCCACTGCGTTTTCCTGCCACAGCATCCAAATCGTGAAGAACTCGGGAGAGG	3059
QY	3430	GTGGAGTCCACATCTAGGGTTGCTGCCCTTGGCTCTATCCCTCCAGAGGTGGAA	3489
Db	3060	GTGGAGTCCACATCTAGGGTTGCTGCCCTTGGCTCTATCCCTCCAGAGGTGGAA	3119
QY	3490	CTGGAGAGTGGGTGCAAGACTGAGCTTAAATGTCTCCCGGCTTGACTTTTCTTCT	3549
Db	3120	CTGGAGAGTGGGTGCAAGACTGAGCTTAAATGTCTCCCGGCTTGACTTTTCTTCT	3179
QY	3550	AGTCTTGGGCTAGATTTCTGCACTTGGGCTCTCTGACAAACACACCAATCCCAAGTAG	3609

Db 3180 AGTCTCTGGGCGCTAGATTCTGCACTTGGGGTCTCTGACACACACACATCCCAAAGTAG 3239  
 QY 3610 CCGGAAGAGCTAAACACAGAGGGGTTCTTAAATGGCTGCCCGCCACCCGGGCTCCCT 3669  
 Db 3240 CCGGAAGAGCTAAACACAGAGGGGTTCTTAAATGGCTGCCCGCCACCCGGGCTCCCT 3299  
 QY 3670 TGGCAAAAGAAATGTGAGCCCTTACCCCAACCCCTTCAACTTACAGAAATCTGGGCCACCC 3729  
 Db 3300 TGGCAAAAGAAATGTGAGCCCTTACCCCAACCCCTTCAACTTACAGAAATCTGGGCCACCC 3359  
 QY 3730 CAGCAGTATTTTATTTAAATGTTGGCCATTTATGAGTTATGATCAATTTGATTAAA 3789  
 Db 3360 CAGCAGTATTTTATTTAAATGTTGGCCATTTATGAGTTATGATCAATTTGATTAAA 3419  
 QY 3790 TTTAAGTTACAGATGTC 3806  
 Db 3420 TTTAAGTTACAGATGTC 3436

RESULT 3  
 ID ADC37268  
 AC ADC37268 standard; DNA; 3243 BP.  
 XX ADC37268;  
 XX 18-DEC-2003 (first entry)  
 XX Nuclear factor kappa B (NF-kappaB) activating gene, SEQ ID 101.  
 XX Nuclear factor kappa B; NF-kappaB, inflammation; autoimmune disease;  
 KW cancer; infectious disease; bone disease; AIDS;  
 KW neurodegenerative disease; ischaemic disorder; Antiinflammatory;  
 KW Immunomodulator; Cytostatic; Antimicrobial; Osteopathic; Anti-HIV;  
 KW Neuroprotective; Nootropic; Cardiant; Gene therapy; human; gene; ds.  
 XX Homo sapiens.  
 OS  
 XX WO2003048202-A2.  
 XX 12-JUN-2003.  
 XX 03-DEC-2002; 2002WO-JP012644.  
 PF  
 XX 03-DEC-2001; 2001JP-00368692.  
 PR 05-DEC-2001; 2001US-0335829P.  
 PR 03-OCT-2002; 2002JP-00291302.  
 PR 04-OCT-2002; 2002US-0415769P.  
 XX (ASAH ) ASahi KASEI KK.  
 PA  
 XX Matsuda A, Muramatsu S;  
 PI  
 XX WPI; 2003-505282/47.  
 DR P-PSDB; ADC37269.  
 DR  
 XX New purified protein that activates nuclear factor kappa B (NF-kappaB),  
 PT useful for treating inflammation, autoimmune diseases, cancers,  
 PT infectious diseases, bone diseases, AIDS, neurodegenerative diseases or  
 PT ischemic disorders.  
 PT  
 XX Claim 4; SEQ ID NO 101; 938pp; English.  
 PS  
 XX The present invention relates to novel proteins and their coding  
 CC sequences (ADC37168-ADC37455), which activate nuclear factor kappa B (NF-  
 CC kappaB). The proteins and their coding sequences are useful for treating  
 CC a disease associated with NF-kappaB activation, such as inflammation,  
 CC autoimmune diseases, cancers, infectious diseases, bone diseases, AIDS,  
 CC neurodegenerative diseases, or ischaemic disorders.  
 XX  
 SQ Sequence 3243 BP; 715 A; 1023 C; 863 G; 642 T; 0 U; 0 Other;

Query Match 78.3%; Score 2994; DB 9; Length 3243;  
 Best Local Similarity 99.9%; Pred. No. 0;

Matches 3234; Conservative 0; Mismatches 0; Indels 2; Gaps 2;  
 QY 571 CTTGGGGCCCCCTCTCTGCTGGCCCTGCGGCGCATGCGCTGACGCTCAAGAGACGAGCTG 630  
 Db 10 CTTGGGGCCCCCTCTCTGCTGGCCCTGCGGCGCATGCGCTGACGCTCAAGAGACGAGCTG 69  
 QY 631 CTGTGCTCCATCTCTGCTGAGCATCTACAGAGACCCGCTGAGCTTGGGCTGCGAGCACTAC 690  
 Db 70 CTGTGCTCCATCTGCTGAGCATCTACAGAGACCCGCTGAGCTTGGGCTGCGAGCACTAC 129  
 QY 691 TTCTGGCGCGCTGCTGCTGAGAGCACTGGGTGGGCGAGAGGCGGAGGGGCGCCCGGAC 750  
 Db 130 TTCTGGCGCGCTGCTGCTGAGAGCACTGGGTGGGCGAGAGGCGGAGGGGCGCCCGGAC 189  
 QY 751 TGGCCGAGTGGCGGCGCAGCTTGGCGGAGCCCGCTGGCGGCGCAGCTCAAGCTGGCC 810  
 Db 190 TGGCCGAGTGGCGGCGCAGCTTGGCGGAGCCCGCTGGCGGCGCAGCTCAAGCTGGCC 249  
 QY 811 AACATCGTGGAGCGCTACAGCTCTTCCGCTGAGCGCCATCTCTCAACGCGCGCGCGCC 870  
 Db 250 AACATCGTGGAGCGCTACAGCTCTTCCGCTGAGCGCCATCTCTCAACGCGCGCGCGCC 309  
 QY 871 GCGGACCTCCAGGCGCAGCAAGGTCAAGCTCTTCTGCTCAGCGACCGCGCGCTT 930  
 Db 310 GCGGACCTCCAGGCGCAGCAAGGTCAAGCTCTTCTGCTCAGCGACCGCGCGCTT 369  
 QY 931 CTCTGCTTCTTCTGCGAGCGCTGCTGCTGAGCGAGCATCAGGTCAACGCGCATCGAC 990  
 Db 370 CTCTGCTTCTTCTGCGAGCGCTGCTGCTGAGCGAGCATCAGGTCAACGCGCATCGAC 429  
 QY 991 GAGCGCTTCAAGCGAGCTGAGAGGAGCTGAAGACCACTTCAAGGCGCTTCAAGACAGC 1050  
 Db 430 GAGCGCTTCAAGCGAGCTGAGAGGAGCTGAAGACCACTTCAAGGCGCTTCAAGACAGC 489  
 QY 1051 GAGCGGAAACACACCGAGCGCTGCTGCTGAGCGCACTTGGCGGAGCACTCAAGTCT 1110  
 Db 490 GAGCGGAAACACACCGAGCGCTGCTGCTGAGCGCACTTGGCGGAGCACTCAAGTCT 549  
 QY 1111 TCACCAAGAGCGCTGCGGACCACTATCGCGAGCGCTTTCAGCGGCTGCAACCGGCTGCTG 1170  
 Db 550 TCACCAAGAGCGCTGCGGACCACTATCGCGAGCGCTTTCAGCGGCTGCAACCGGCTGCTG 609  
 QY 1171 CGTGAACCGCAGAGGCCATGCTAGAGAGCTGAGAGCGGACACGCGCCGACGCTGACC 1230  
 Db 610 CGTGAACCGCAGAGGCCATGCTAGAGAGCTGAGAGCGGACACGCGCCGACGCTGACC 669  
 QY 1231 GACATCGAGCAGAAAGTCCAGCGCTACAGCAGCGCTGCGCAGAGCTCCAGAGGAGGCC 1290  
 Db 670 GACATCGAGCAGAAAGTCCAGCGCTACAGCAGCGCTGCGCAGAGCTCCAGAGGAGGCC 729  
 QY 1291 CAGATCTCGAGGAGCGGCTGGAACCGACCGCGCACACTTCTCTGGCTGGGCTGGCC 1350  
 Db 730 CAGATCTCGAGGAGCGGCTGGAACCGACCGCGCACACTTCTCTGGCTGGGCTGGCC 789  
 QY 1351 TCACTGTCGAGCGGCTCAAGGGAAAAATCCATGAGACCACTCAATATGAAGACTTC 1410  
 Db 790 TCACTGTCGAGCGGCTCAAGGGAAAAATCCATGAGACCACTCAATATGAAGACTTC 849  
 QY 1411 CCGACCTCCAGTACAGAGCGGCTGCTGCTGAGTACACCATCTGGAAGTCCCTGTTCCAGGAC 1470  
 Db 850 CCGACCTCCAGTACAGAGCGGCTGCTGCTGAGTACACCATCTGGAAGTCCCTGTTCCAGGAC 909  
 QY 1471 ATCCACCACTGCGCAGCGGCTTAACCTGGAGCCCGGGGACAGCCCAAGCGGCTGATC 1530  
 Db 910 ATCCACCACTGCGCAGCGGCTTAACCTGGAGCCCGGGGACAGCCCAAGCGGCTGATC 969  
 QY 1531 CTGTGGAACAGCTGACCATTTGGGCTTACGGCACTTGCACCCACAGCCACTTCAGGAC 1590  
 Db 970 CTGTGGAACAGCTGACCATTTGGGCTTACGGCACTTGCACCCACAGCCACTTCAGGAC 1029  
 QY 1591 TCGCCAAAGCGCTTCGATGTTGGAGGCTGCTGGTCTGAGGCTTCTGAGGCTTCAGTAGTGGC 1650  
 Db 1030 TCGCCAAAGCGCTTCGATGTTGGAGGCTGCTGGTCTGAGGCTTCTGAGGCTTCAGTAGTGGC 1089



1651 GTCCACTACTGGAGGTGGTGGGAGAGAACCCAGTGGGTGATCGGGCTGGCACAC 1710  
 1090 GTCCACTACTGGAGGTGGTGGGAGAGAACCCAGTGGGTGATCGGGCTGGCACAC 1149  
 1711 GAAGCGCAAGCGCAAGCGCAAGCGCAAGCGCAAGCGCAAGCGCAAGCGCAAGCGCA 1770  
 1150 GAAGCGCAAGCGCAAGCGCAAGCGCAAGCGCAAGCGCAAGCGCAAGCGCAAGCGCA 1209  
 1771 GTGATGACGATGGCAACACAGTACAGCGCTGACGAGGCGCTGACGAGCGGCTTAACGTC 1830  
 1210 GTGATGACGATGGCAACACAGTACAGCGCTGACGAGGCGCTGACGAGCGGCTTAACGTC 1269  
 1831 CGGGCAAGCTTGACAAAGTGGGTGCTTCCCTGGACTATGACCAAGGCTTGCTCATCTTC 1890  
 1270 CGGGCAAGCTTGACAAAGTGGGTGCTTCCCTGGACTATGACCAAGGCTTGCTCATCTTC 1329  
 1891 TACAATGCTGATGACATGCTCTGGCTTACACCTCCGCGAGAAGTTCCCTGGCAAGCTC 1950  
 1330 TACAATGCTGATGACATGCTCTGGCTTACACCTCCGCGAGAAGTTCCCTGGCAAGCTC 1389  
 1951 TGCTCTTACTTACGCTCGCAGAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAG 2010  
 1390 TGCTCTTACTTACGCTCGCAGAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAG 1449  
 2011 ATCAACACCGTCCGATCTAGTCCAGGCAAGAGGACACCAACCTCTCTGGGACCACTGC 2070  
 1450 ATCAACACCGTCCGATCTAGTCCAGGCAAGAGGACACCAACCTCTCTGGGACCACTGC 1509  
 2071 CACCTGCAAGAGCCCTGCGCAGGAGATAGAGACCTGGACTCCAGCCACCGTGGCCAC 2130  
 1510 CACCTGCAAGAGCCCTGCGCAGG-AGATAGAGACCTGGACTCCAGCCACCGTGGCCAC 1568  
 2131 TGGAGACCTCAGGACAGTTGTTTACCTCCAGCTCCAGTCTGTGTAAATGGAGGTGGAT 2190  
 1569 TGGAGACCTCAGGACAGTTGTTTACCTCCAGCTCCAGTCTGTGTAAATGGAGGTGGAT 1628  
 2191 TCCCTACTTCCCTAACTCTTCCAGCATCGATGTTCTGTAGCTCTGACTGTGATAGGGA 2250  
 1629 TCCCTACTTCCCTAACTCTTCCAGCATCGATGTTCTGTAGCTCTGACTGTGATAGGGA 1688  
 2251 TACAGCTTTGATCCAGGATGTGACATGGCTTCTCCTCAGGCAACCCCTGCGCCACCCCT 2310  
 1689 TACAGCTTTGATCCAGGATGTGACATGGCTTCTCCTCAGGCAACCCCTGCGCCACCCCT 1748  
 2311 CATCCCCATCTTCTCAGGCGCAGGGGACTACCTTCCAGTGTCTCCCTCAGCCAGCCCT 2370  
 1749 CATCCCCATCTTCTCAGGCGCAGGGGACTACCTTCCAGTGTCTCCCTCAGCCAGCCCT 1808  
 2371 GACCTCAGGAAGTGTGACAGCATGGCCAGTGTGGCAGCCGGAAGACACACAGCACCC 2430  
 1809 GACCTCAGGAAGTGTGACAGCATGGCCAGTGTGGCAGCCGGAAGACACACAGCA-CC 1867  
 2431 TCTTATGTCCTATGGCTTAAGACTTACCCCTGACCAAGCTAGTGTGGGCCCAATTTACCT 2490  
 1868 TCTTATGTCCTATGGCTTAAGACTTACCCCTGACCAAGCTAGTGTGGGCCCAATTTACCT 1927  
 2491 TGACCCCACTCAGAGTGTGACAGTAGTACCTGTCTCCTAGGTTGCTCCTGAGAGCCAAC 2550  
 1928 TGACCCCACTCAGAGTGTGACAGTAGTACCTGTCTCCTAGGTTGCTCCTGAGAGCCAAC 1987  
 2551 CTCTCTCTGCCCCACACCAAGAACTATATGTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2610  
 1988 CTCTCTCTGCCCCACACCAAGAACTATATGTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2047  
 2611 GTCAAGTATGATGCTGTGGCTGTGAGGAGCACTGGTAGTTGAGTCCACACATTATAGT 2670  
 2048 GTCAAGTATGATGCTGTGGCTGTGAGGAGCACTGGTAGTTGAGTCCACACATTATAGT 2107  
 2671 CATGTGCCACACCTTCTCTGCCCCACAGGCCGAGGAGCAAGGTTGAGGTTATACCCAAAGCT 2730  
 2108 CATGTGCCACACCTTCTCTGCCCCACAGGCCGAGGAGCAAGGTTGAGGTTATACCCAAAGCT 2167

2731 GATCAGAGCCCATTTAGCCTAAAAAGCAACTGCAGGACAAAGCCCTCCTCGATGATCGAGGT 2790  
 2168 GATCAGAGCCCATTTAGCCTAAAAAGCAACTGCAGGACAAAGCCCTCCTCGATGATCGAGGT 2227  
 2791 CCCCAGTAGCTCTGAACAAGAGTCCAGCCAAACCCCTTTCAGCCAGGCGCTCTGTGACCTGC 2850  
 2228 CCCCAGTAGCTCTGAACAAGAGTCCAGCCAAACCCCTTTCAGCCAGGCGCTCTGTGACCTGC 2287  
 2851 TAGGGTGCAGGAGGCTTCCAGAGCAAGTGTGTTGTAATTAGGACCCCAAGCACTGGGAGGG 2910  
 2288 TAGGGTGCAGGAGGCTTCCAGAGCAAGTGTGTTGTAATTAGGACCCCAAGCACTGGGAGGG 2347  
 2911 CTGTTGGCTAGACCCCTTCTCAGACTTGGCACTATCTCAGTTAGGATCTCTGCTGCAGAA 2970  
 2348 CTGTTGGCTAGACCCCTTCTCAGACTTGGCACTATCTCAGTTAGGATCTCTGCTGCAGAA 2407  
 2971 AACAGAGCCACTTGTAGCTGTTTAAATPAGAACAGGATTTACTACCTGGCCCCCTGGTGG 3030  
 2408 AACAGAGCCACTTGTAGCTGTTTAAATPAGAACAGGATTTACTACCTGGCCCCCTGGTGG 2467  
 3031 CTTGCAGAAATTTGTTGAGAGCTGGAGAGCAGACTCTCTGTAATTTCCAGAACTCCCA 3090  
 2468 CTTGCAGAAATTTGTTGAGAGCTGGAGAGCAGACTCTCTGTAATTTCCAGAACTCCCA 2527  
 3091 GCGCAGATTCATCATCTGTCTGTGTGACCAAGAAAGCTGCCCCATCTCAGAGAAAGCCAC 3150  
 2528 GCGCAGATTCATCATCTGTGTGTGACCAAGAAAGCTGCCCCATCTCAGAGAAAGCCAC 2587  
 3151 TATGCCAGAAAGCTGTGACTGCAGAACTAGCTCCCTCTGCCAGGTCCTGGCCAGCCCA 3210  
 2588 TATGCCAGAAAGCTGTGACTGCAGAACTAGCTCCCTCTGCCAGGTCCTGGCCAGCCCA 2647  
 3211 ATAGATGCTCAGGCGCTGCCCTCTCCACTCTCAGTCTCAGTCTCCCAAACTTAAATTTTA 3270  
 2648 ATAGATGCTCAGGCGCTGCCCTCTCCACTCTCAGTCTCAGTCTCCCAAACTTAAATTTTA 2707  
 3271 CAAGAGATTCGTGTTGGGGAACTTAAAGTCAGATCCAGAACTTGGCTCCAGAGGAGTCT 3330  
 2708 CAAGAGATTCGTGTTGGGGAACTTAAAGTCAGATCCAGAACTTGGCTCCAGAGGAGTCT 2767  
 3331 GGGAAATGTCATTTCCCTAGAAAGGATTTAGGTTGGTGGTGGAGCAAGCCCCACCTCGTTT 3390  
 2768 GGGAAATGTCATTTCCCTAGAAAGGATTTAGGTTGGTGGTGGAGCAAGCCCCACCTCGTTT 2827  
 3391 TTCTGCCACAGCATCCAAATCGTAGAAGTCCGGAGAGGTTGGAGTCCACATCTAGGTT 3450  
 2828 TTCTGCCACAGCATCCAAATCGTAGAAGTCCGGAGAGGTTGGAGTCCACATCTAGGTT 2887  
 3451 GTCTTCCCTTGGCTCTATCCCTGCCAGAGTGGGAACTGGAGAGTGGGCTGCAAGA 3510  
 2888 GTCTTCCCTTGGCTCTATCCCTGCCAGAGTGGGAACTGGAGAGTGGGCTGCAAGA 2947  
 3511 CTGAGCTTAAATGTCTCCCGGCTTGAATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3570  
 2948 CTGAGCTTAAATGTCTCCCGGCTTGAATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3007  
 3571 CACTTGGGCTCTCTGACACACACACCATCCAAAGTAGCCGGAAGAGCTAAACACAGG 3630  
 3008 CACTTGGGCTCTCTGACACACACACCATCCAAAGTAGCCGGAAGAGCTAAACACAGG 3067  
 3631 GGTCTTAAATGGTGGCCCCCGCCAGCCGCGCTCCCTTGGGCAAAAGAAATTTGTACG 3690  
 3068 GGTCTTAAATGGTGGCCCCCGCCAGCCGCGCTCCCTTGGGCAAAAGAAATTTGTACG 3127  
 3691 CTTACCCCAACCCCTTCAACTACAGAACTCTGGGCCACCCCAAGCAGTATTTTATTTAAA 3750  
 3128 CTTACCCCAACCCCTTCAACTACAGAACTCTGGGCCACCCCAAGCAGTATTTTATTTAAA 3187  
 3751 TGTTCGCCCAATTTATGAGTTATGATCAATTTGTATTAAATTAAGTTACAGATGTC 3806  
 3188 TGTTCGCCCAATTTATGAGTTATGATCAATTTGTATTAAATTAAGTTACAGATGTC 3243



AC ABQ92018;  
 XX  
 DT 04-OCT-2002 (first entry)  
 XX  
 DE Human polynucleotide SEQ ID NO 15.  
 XX  
 KW Human; cytostatic; antirheumatic; antiarthritic; vulnery; analgesic;  
 KW antiinflammatory; antibacterial; immunosuppressive; antiparkinsonian;  
 KW neuroprotective; nootropic; osteopathic; haemostatic; vasotropic;  
 KW antitumor; fungicide; antidiabetic; antiasthmatic; antiallergic;  
 KW immunostimulant; antiparasitic; secreted protein; transmembrane protein;  
 KW cytokine; cell proliferation; cell differentiation; autoimmune disease;  
 KW stem cell; growth factor; nervous system disease; neuropathy;  
 KW Alzheimer's disease; Parkinson's disease; Huntington's disease;  
 KW osteoporosis; severe combined immunodeficiency; SCID; infection;  
 KW multiple sclerosis; rheumatoid arthritis; gene therapy; gene; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 XX US2002065394-A1.  
 XX  
 XX 30-MAY-2002.  
 XX  
 XX 22-DEC-2000; 2000US-00745763.  
 XX  
 XX 18-MAR-1998; 98US-00040963.  
 XX  
 XX (JACO/) MCCOY J M.  
 XX (MCCO/) MCCOY J M.  
 XX (LAVA/) LAVALLIE E R.  
 XX (COLL/) COLLINS-RACIE L A.  
 XX (EVAN/) EVANS C.  
 XX (MERB/) MERBERG D.  
 XX (TREAC/) TREACY M.  
 XX (SPAUL/) SPAULDING V.  
 XX  
 XX Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;  
 XX Merberg D, Treacy M, Spaulding V;  
 XX  
 XX WPI: 2002-582343/62.  
 XX P-PSDB; ABP61802.  
 XX  
 XX Novel secreted or transmembrane protein and polynucleotide encoding the  
 XX protein, useful for diagnosis and treatment of neurological disorders.  
 XX Cancer, autoimmune diseases, bone disorders and lung or liver fibrosis.  
 XX  
 XX Claim 56; Page 117-118; 284pp; English.  
 XX  
 XX The invention relates to human secreted or transmembrane protein (I),  
 XX their fragments and is encoded by specific complementary deoxyribonucleic  
 XX acid (cDNA) inserts (II), where the protein is substantially free from  
 XX other mammalian proteins. (I) are useful for preventing, treating or  
 XX ameliorating a medical condition, especially immunological treatment or  
 XX prevention of tumours. (I) exhibits activity relating to angiogenesis,  
 XX cytokine, cell proliferation, cell differentiation, antiinflammatory,  
 XX stem cell growth factor activity and activin or inhibin-related  
 XX activities. (I) can be used to manipulate stem cells in culture to give  
 XX rise to neuroepithelial cells that can be used to augment or replace  
 XX cells damaged by illness, autoimmune disease, accidental damage or  
 XX genetic disorders. (I) induces the proliferation of neural cells and  
 XX regeneration of nerve and brain tissue and is useful for the treatment of  
 XX central and peripheral nervous system diseases and neuropathies, such as  
 XX Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
 XX lateral sclerosis. (I) is involved in chemotactic or chemokinetic  
 XX activity, regulation of haematopoiesis and is useful for treating myeloid  
 XX or lymphoid cell disorders, platelet disorders such as thrombocytopaenia  
 XX and for regeneration of bone, cartilage, tendon, ligament and/or nerve  
 XX tissue growth and in tissue repair, healing of burns, incisions, ulcers,  
 XX for treating osteoporosis, osteoarthritis, bone degenerative disorders or  
 XX periodontal disease. (I) is also useful for gut protection or  
 XX regeneration and treatment of lung or liver fibrosis, reperfusion injury  
 XX in various tissues, various immune deficiencies and disorders including  
 XX severe combined immunodeficiency (SCID), bacterial or fungal infections,

CC autoimmune disorders e.g. multiple sclerosis, rheumatoid arthritis,  
 CC diabetes mellitus, myasthenia gravis, allergic reactions and conditions,  
 CC such as asthma or other respiratory problems. (II) is useful to express  
 CC recombinant protein, as markers for tissues in which the corresponding  
 CC protein is preferentially expressed and in gene therapy. The present  
 CC sequence is that of a polynucleotide of the invention  
 XX  
 XX Sequence 2522 BP; 558 A; 683 C; 608 G; 673 T; 0 U; 0 Other;  
 Query Match 0.7%; Score 25; DB 6; Length 2522;  
 Best Local Similarity 100.0%; Pred. No. 28;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3802 ATGTCAAAAAAAAAAAAAAAAAAAAAA 3826  
 DB 2495 ATGTCAAAAAAAAAAAAAAAAAAAAAA 2519  
 RESULT 7  
 ADB62721/c  
 ID ADB62721 standard; cDNA; 2586 BP.  
 XX  
 AC ADB62721;  
 XX  
 XX 04-DEC-2003 (first entry)  
 XX  
 XX Human cDNA encoding clone NT2NE20028700.  
 XX  
 XX Human; ss; gene; pharmaceutical; diagnostic; gene therapy;  
 XX tissue regeneration; cell regeneration; membrane protein;  
 XX signal transduction-related protein; transcription-related protein;  
 XX osteoporosis; neurological disease; cancer; tumour.  
 XX  
 XX Homo sapiens.  
 XX  
 XX Key Location/Qualifiers  
 XX CDS 166..1377  
 XX FT /\*tag= a  
 XX FT /product= "Clone NT2NE20028700 protein"  
 XX  
 XX EP1308459-A2.  
 XX  
 XX 07-MAY-2003.  
 XX  
 XX 28-MAR-2002; 2002EP-00007401.  
 XX  
 XX 05-NOV-2001; 2001JP-00379298.  
 XX 25-JAN-2002; 2002US-00350978.  
 XX  
 XX (HELI-) HELIX RES INST.  
 XX (REAS-) RES ASSOC BIOTECHNOLOGY.  
 XX  
 XX Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;  
 XX Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;  
 XX Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;  
 XX WPI: 2003-450961/43.  
 XX P-PSDB; ADB64691.  
 XX  
 XX New polynucleotides and polypeptides, useful for developing a diagnostic  
 XX marker or medicines for regulation of their expression and activity, or  
 XX as targets of gene therapy.  
 XX  
 XX Claim 1; Page; 222pp; English.  
 XX  
 XX The invention discloses a polynucleotide comprising a sequence selected  
 XX from 1970 fully defined nucleotide sequences which encode novel  
 XX polypeptides. Also claimed is a polypeptide encoded by the polynucleotide  
 XX or its partial peptide, an antibody binding to the polypeptide or peptide  
 XX of the polynucleotide, immunologically assaying the polypeptide or  
 XX peptide of the polynucleotide by contacting the polypeptide or peptide  
 XX with the antibody of the encoded protein, and observing the binding  
 XX between the two, a transformant carrying the polynucleotide in an



SQ Sequence 2647 BP; 942 A; 475 C; 384 G; 837 T; 0 U; 9 Other;  
Query Match 0.7%; Score 25; DB 5; Length 2647;  
Best Local Similarity 100.0%; Pred. No. 28;  
Matches 25; Conservative 0; Mismatches 0; Indels

Qy 3802 ATGTCAAAAAAAAAAAAAAAAAAAA 3826  
|||  
Db 58 ATGTCAAAAAAAAAAAAAAAAAAAA 34

RESULT 10	
ABV25443/c	
ID	ABV25449 standard; cDNA; 2647 BP.
XX	
XX	
AC	ABV25449;
XX	
XX	
DT	16-SEP-2002 (first entry)
XX	
DE	
XX	
XX	Human prostate expression marker cDNA 25440.
XX	
XX	Human; prostate cancer; cytostatic; carcinog
KW	pharmacogenomic marker; gene; ss
KW	

Query Match 0.7%; Score 25; DB 5; Length 2647;  
Best Local Similarity 100.0%; Pred.No.28;  
Matches 25; Conservative 0; Mismatches 0; Indels

Qy 3802 ATGTCAAAAAAAAAAAAAAAAAAAA 3826  
|||  
Db 58 ATGTCAAAAAAAAAAAAAAAAAAAA 34

RESULT 11	
ABV30164/c	
ID	ABV30164 standard; cDNA; 2647 BP.
XX	
XX	
AC	ABV30164;
XX	
XX	
DT	16-SEP-2002 (first entry)
XX	
DE	Human prostate expression marker cDNA 30155.
XX	
XX	
KW	Human; prostate cancer; cytostatic; carcinog
KW	pharmacogenomic marker; gene; ss

Query Match 0.7%; Score 25; DB 5; Length 2647;  
Best Local Similarity 100.0%; Pred. No. 28;  
Matches 25; Conservative 0; Mismatches 0; Indels

QY 3802 ATGTCAAAAAAAAAAAAAAAAAAAA 3826  
|||  
Db 58 ATGTCAAAAAAAAAAAAAAAAAAAA 34

RESULT 12  
ABV25571/c  
ID ABV25571 standard; cDNA; 2647 BP.

DT 16-SEP-2002 (first entry)

```
XX Human prostate expression marker cDNA 25562.
DE
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX Homo sapiens.
XX WO200160860-A2.
XX 23-AUG-2001.
XX PF
XX 20-FEB-2001; 2001WO-US005171.
XX PI
XX 17-FEB-2000; 2000US-0183319P.
XX 16-MAR-2000; 2000US-0189862P.
XX 25-MAY-2000; 2000US-0207454P.
XX 09-JUN-2000; 2000US-0211314P.
XX 18-JUL-2000; 2000US-0219007P.
XX 13-DEC-2000; 2000US-0255281P.
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX PI
XX Schlegel R, Endege WO, Monahan JE;
XX WPI; 2001-662795/76.
XX Novel isolated nucleic acid molecule associated with cancerous state of
XX prostate cells and correlating with presence of prostate cancer, useful
XX for detecting presence of prostate cancer, stage of prostate cancer.
XX Claim 1; Page 5100; 11750pp; English.
XX The invention relates to an isolated nucleic acid molecule (I) comprising
XX a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
XX specification or its complement. (I) is useful for: (a) assessing whether
XX a patient is afflicted with prostate cancer; (b) monitoring the
XX progression of prostate cancer in a patient; (c) assessing the efficacy
XX of a test compound to inhibit prostate cancer in a patient; (d) assessing
XX the efficacy of a therapy for inhibiting prostate cancer in a patient;
XX (e) selecting a composition for inhibiting prostate cancer in a patient;
XX (f) assessing the prostate cell carcinogenic potential of a compound; (g)
XX determining whether prostate cancer has metastasized in a patient; (h)
XX assessing the aggressiveness or indolence of prostate cancer in a patient
XX ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
XX Sequence 2647 BP; 942 A; 475 C; 384 G; 837 T; 0 U; 9 Other;
XX
XX Query Match 0.7%; Score 25; DB 5; Length 2647;
XX Best Local Similarity 100.0%; Pred. No. 28;
XX Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 3802 ATGTCAAAAA...AAAAAAAAA 3826
XX DB 58 ATGTCAAAAA...AAAAAAAAA 34
XX
XX Sequence 2647 BP; 942 A; 475 C; 384 G; 837 T; 0 U; 9 Other;
XX
XX Query Match 0.7%; Score 25; DB 5; Length 2647;
XX Best Local Similarity 100.0%; Pred. No. 28;
XX Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 3802 ATGTCAAAAA...AAAAAAAAA 3826
XX DB 58 ATGTCAAAAA...AAAAAAAAA 34
XX
XX RESULT 13
XX ABV29143/C
XX ID ABV29143 standard; cDNA; 2647 BP.
XX XX
XX AC ABV29143;
XX XX
XX 16-SEP-2002 (first entry)
XX XX
XX Human prostate expression marker cDNA 29134.
XX DE
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX KW pharmacogenomic marker; gene; ss.
XX OS Homo sapiens.
XX XX
XX WO200160860-A2.
XX PN
```

```
XX 23-AUG-2001.
XX 20-FEB-2001; 2001WO-US005171.
XX 17-FEB-2000; 2000US-0183319P.
XX 16-MAR-2000; 2000US-0189862P.
XX 25-MAY-2000; 2000US-0207454P.
XX 09-JUN-2000; 2000US-0211314P.
XX 18-JUL-2000; 2000US-0219007P.
XX 13-DEC-2000; 2000US-0255281P.
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX PI
XX Schlegel R, Endege WO, Monahan JE;
XX WPI; 2001-662795/76.
XX Novel isolated nucleic acid molecule associated with cancerous state of
XX prostate cells and correlating with presence of prostate cancer, useful
XX for detecting presence of prostate cancer, stage of prostate cancer.
XX Claim 1; Page 6196-6197; 11750pp; English.
XX The invention relates to an isolated nucleic acid molecule (I) comprising
XX a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
XX specification or its complement. (I) is useful for: (a) assessing whether
XX a patient is afflicted with prostate cancer; (b) monitoring the
XX progression of prostate cancer in a patient; (c) assessing the efficacy
XX of a test compound to inhibit prostate cancer in a patient; (d) assessing
XX the efficacy of a therapy for inhibiting prostate cancer in a patient;
XX (e) selecting a composition for inhibiting prostate cancer in a patient;
XX (f) assessing the prostate cell carcinogenic potential of a compound; (g)
XX determining whether prostate cancer has metastasized in a patient; (h)
XX assessing the aggressiveness or indolence of prostate cancer in a patient
XX ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
XX Sequence 2647 BP; 942 A; 475 C; 384 G; 837 T; 0 U; 9 Other;
XX
XX Query Match 0.7%; Score 25; DB 5; Length 2647;
XX Best Local Similarity 100.0%; Pred. No. 28;
XX Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 3802 ATGTCAAAAA...AAAAAAAAA 3826
XX DB 58 ATGTCAAAAA...AAAAAAAAA 34
XX
XX Sequence 2647 BP; 942 A; 475 C; 384 G; 837 T; 0 U; 9 Other;
XX
XX Query Match 0.7%; Score 25; DB 5; Length 2647;
XX Best Local Similarity 100.0%; Pred. No. 28;
XX Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 3802 ATGTCAAAAA...AAAAAAAAA 3826
XX DB 58 ATGTCAAAAA...AAAAAAAAA 34
XX
XX RESULT 14
XX AAX35727
XX ID AAX35727 standard; cDNA; 2796 BP.
XX XX
XX AC AAX35727;
XX XX
XX 09-JUL-1999 (first entry)
XX XX
XX cDNA encoding a protein identified by the signal sequence trap method.
XX
XX Signal sequence trap method, SST method; immunisation; inhibition;
XX infection; allergy; cancer; regulation; tissue formation; tissue repair;
XX activin activity; inhibin activity; chemokine activity;
XX cytokine activity; blood coagulation regulation; agonist; antagonist;
XX metabolic disorder; hormonal disorder; immune disorder;
XX severe combined immunodeficiency; SCID; AIDS; thrombosis; cancer; wound;
XX ss.
XX XX
XX Homo sapiens.
XX OS
XX WO9918126-A1.
XX PN
XX 15-APR-1999.
XX PD
XX 06-OCT-1998; 98WO-JP004514.
XX PF
```

```

XX 07-OCT-1997; 97JP-00274674.
XX (ONCOY ) ONO PHARM CO LTD.
XX
XX Fukushima D, Shibayama S, Tada H;
XX WPI; 1999-277254/23.
XX DR P-PSDB; AAY02374.
XX
XX Polypeptides identified by the signal sequence trap method from a human
XX cDNA library.
XX
XX Claim 4; Page 153-157; 281pp; Japanese.
XX
XX AAX35694-X35747 represent cDNA sequences that encode novel polypeptides
XX (AAY02358-84) which are identified from a human placental cDNA library by
XX the signal sequence trap (SST) method. The polypeptides have a broad
XX range of physiological activity, including immunisation against and
XX inhibition of infectious, allergies and cancer; regulation of tissue
XX formation and repair; activin/inhibin activity; chemokine/cytokine
XX activity; blood coagulation regulation; and receptor/ligand agonist or
XX antagonist activity. The polypeptides can be used for prevention and
XX treatment of disorders including infections by bacteria, yeasts and
XX viruses (including HIV) and protozoa; metabolic and hormonal disorders;
XX immune disorders (including severe combined immunodeficiency (SCID) and
XX AIDS; thrombosis; cancer; and traumatic or surgical wounds
XX
XX Sequence 2796 BP; 636 A; 690 C; 674 G; 796 T; 0 U; 0 Other;
XX
XX Query Match 0.7%; Score 25; DB 2; Length 2796;
XX Best Local Similarity 100.0%; Pred. No. 28;
XX Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 3802 ATGTCAAAAAAAAAAAAAAAAAAAAA 3826
Db |||||||||||||||||||||||||||
2768 ATGTCAAAAAAAAAAAAAAAAAAAAA 2792

RESULT 15
ADD69641
ID ADD69641 standard; cDNA; 3858 BP.
XX
XX AC ADD69641;
XX
XX 15-JAN-2004 (first entry)
XX
XX Human REMAP cDNA - SEQ ID 70.
XX
XX human; receptor and membrane-associated protein; REMAP; cytoskeletal;
XX antiarteriosclerotic; anticonvulsant; neurotropic; neuroprotective;
XX cerebroprotective; anti-HIV; anti-allergic; anti-inflammatory;
XX thymomimetic; cell proliferative; cancer; atherosclerosis; neurological;
XX epilepsy; Huntington's disease; stroke; immune; inflammatory; AIDS;
XX allergy; developmental; hypothyroidism; Cushing's syndrome; infection;
XX ss; gene.
XX
XX Homo sapiens.
XX
XX OS
XX
XX WO2003048305-A2.
XX
XX 12-JUN-2003.
XX
XX 13-NOV-2002; 2002WO-US036759.
XX
XX 13-NOV-2001; 2001US-0333097P.
XX 15-NOV-2001; 2001US-0335274P.
XX 14-DEC-2001; 2001US-0340542P.
XX 18-DEC-2001; 2001US-0342166P.
XX 11-JAN-2002; 2002US-0347580P.
XX 14-JAN-2002; 2002US-0348687P.
XX
XX (INCY-) INCYTE GENOMICS INC.

```

```

XX Duggan BM, Yang J, Gietzen KJ, Lee SY, Tang YT, Azimzai Y;
XX Chawla NK, Warren BA, Barroso I, Becha SD, Yue H, Lehr-Mason PM;
XX Thangavelu K, Lee S, Emerling BM, Kadle AE, Khare R, Baughn MK;
XX Gandhi AR, Tran UK, Richardson TW, Marquis JP, Lal PG, Forsythe IJ;
XX Lee EA, Swarnakar A, Kallick DA, Griffin JA, Elliott VS, Gorvad AE;
XX Hafalia AJA, Ison CH, Jin P, Jiang X, Jackson AA, Bhatia U;
XX Burrill JD, Blake JJ, Ho A, Zheng W, Gao J;
XX
XX WPI; 2003-513744/48.
XX DR P-PSDB; ADD69594.
XX
XX New human receptors and membrane-associated proteins (REMAP), useful for
XX diagnosing, treating and preventing diseases or conditions associated
XX with the aberrant REMAP expression e.g. cancer, AIDS, atherosclerosis, or
XX infections.
XX
XX Claim 5; SEQ ID NO 70; 298pp; English.
XX
XX The invention relates to a novel isolated polypeptide comprising a human
XX receptor and membrane-associated protein (REMAP). The polypeptide of the
XX invention demonstrates cytostatic, antiarteriosclerotic, anticonvulsant,
XX neurotropic, neuroprotective, cerebroprotective, anti-HIV, anti-allergic,
XX anti-inflammatory and thymomimetic activities and may be useful for
XX treating and diagnosing various disorders including those which are cell
XX proliferative such as cancer and atherosclerosis, neurological including
XX epilepsy, Huntington's disease and stroke, immune/inflammatory
XX particularly AIDS and allergies and developmental such as hypothyroidism
XX and Cushing's syndrome, as well as infections. The current sequence is
XX that of the human REMAP cDNA of the invention.
XX
XX Sequence 3858 BP; 972 A; 961 C; 989 G; 936 T; 0 U; 0 Other;
XX
XX Query Match 0.7%; Score 25; DB 9; Length 3858;
XX Best Local Similarity 100.0%; Pred. No. 26;
XX Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 3802 ATGTCAAAAAAAAAAAAAAAAAAAAA 3826
Db |||||||||||||||||||||||||||
3815 ATGTCAAAAAAAAAAAAAAAAAAAAA 3839

Search completed: September 4, 2004, 08:48:49
Job time : 950 secs

```



**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 4, 2004, 08:25:37 ; Search time 183 Seconds  
(without alignments)  
11602.416 Million cell updates/sec

Title: US-09-927-091-3  
Perfect score: 3826  
Sequence: 1 agcgtgcgtgacgcgaacg.....aaaaaaaaaaaaaaaaaaaa 3826

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 682709 seqs, 277475446 residues

Word size : 18

Total number of hits satisfying chosen parameters: 1852

Minimum DB seq length: 2500  
Maximum DB seq length: 5000

Post-processing: Listing first 45 summaries

Database : Issued Patents NA:  
1: /cgn2\_6/prodata/2/ina/5A COMB.seq.\*  
2: /cgn2\_6/prodata/2/ina/5B COMB.seq.\*  
3: /cgn2\_6/prodata/2/ina/6A COMB.seq.\*  
4: /cgn2\_6/prodata/2/ina/6B COMB.seq.\*  
5: /cgn2\_6/prodata/2/ina/PCTUS COMB.seq.\*  
6: /cgn2\_6/prodata/2/ina/backfiles1.seq.\*

pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	24	0.6	3001	4	US-09-539-333D-222
2	24	0.6	3848	3	US-09-112-096-28
3	24	0.6	3952	2	US-08-381-691-16
4	23	0.6	2759	4	US-09-144-367-1
5	23	0.6	2929	4	US-09-705-299-10
6	23	0.6	3238	3	US-08-123-934A-5
7	23	0.6	3238	4	US-09-874-628-5
8	23	0.6	3238	5	PCT-US94-10080-5
9	23	0.6	3268	4	US-09-688-188B-1
10	23	0.6	3268	4	US-09-231-417D-1
11	23	0.6	3300	4	US-09-336-643A-82
12	23	0.6	3377	4	US-09-819-389-1
13	23	0.6	3377	4	US-10-273-992-1
14	23	0.6	3441	4	US-09-866-028-6
15	23	0.6	3946	3	US-09-083-351-1
16	23	0.6	3946	3	US-09-083-352-1
17	23	0.6	4113	4	US-09-785-381-2
18	23	0.6	4843	3	US-08-986-485-1
19	23	0.6	4874	4	US-09-187-330-2
20	22	0.6	2525	4	US-09-229-583A-1
21	22	0.6	2525	4	US-10-187-904-1
22	22	0.6	2544	3	US-09-518-046-3
23	22	0.6	2555	2	US-08-960-022-15
24	22	0.6	2567	3	US-08-993-260-4
25	22	0.6	2575	4	US-09-716-129-25
26	22	0.6	2584	4	US-09-716-129-47
27	22	0.6	2593	4	US-09-614-912-89

28	0.6	2606	4	US-09-234-827B-3	Sequence 3, Appli
29	0.6	2657	4	US-09-495-050A-191	Sequence 191, App
30	0.6	2668	4	US-09-370-838-156	Sequence 156, App
31	0.6	2693	4	US-09-608-285A-48	Sequence 48, Appl
32	0.6	2693	4	US-09-557-800C-48	Sequence 48, Appl
33	0.6	2718	4	US-09-667-135-1	Sequence 1, Appli
34	0.6	2729	1	US-07-938-782A-1	Sequence 1, Appli
35	0.6	2729	1	US-08-630-524-1	Sequence 1, Appli
36	0.6	2729	5	PCT-US93-08131-1	Sequence 1, Appli
37	0.6	2752	4	US-09-636-791A-14	Sequence 14, Appl
38	0.6	2762	4	US-09-608-285A-26	Sequence 26, Appl
39	0.6	2762	4	US-09-608-285A-52	Sequence 52, Appl
40	0.6	2762	4	US-09-240-639-1	Sequence 1, Appli
41	0.6	2762	4	US-09-370-265-26	Sequence 26, Appl
42	0.6	2762	4	US-09-557-800C-26	Sequence 26, Appl
43	0.6	2762	4	US-09-557-800C-52	Sequence 52, Appl
44	0.6	2762	4	US-09-370-625A-26	Sequence 26, Appl
45	0.6	2779	4	US-09-149-476-191	Sequence 191, App

## ALIGNMENTS

RESULT 1  
US-09-539-333D-222  
; Sequence 222, Application US/09539333D  
; Patent No. 6476208  
; GENERAL INFORMATION:  
; APPLICANT: Cohen, Daniel  
; APPLICANT: Blumenfeld, Marta  
; APPLICANT: Chumakov, Ilya  
; APPLICANT: Bougueleret, Lydie  
; APPLICANT: Bihain, Bernard  
; APPLICANT: Essieux, Laurent  
; TITLE OF INVENTION: SCHIZOPHRENIA ASSOCIATED GENES, PROTEINS AND BIALLELIC MARKERS  
; FILE REFERENCE: GENSET.047AUS  
; CURRENT APPLICATION NUMBER: US/09/539,333D  
; CURRENT FILING DATE: 2000-03-30  
; PRIOR APPLICATION NUMBER: US 60/126,903  
; PRIOR FILING DATE: 1999-03-30  
; PRIOR APPLICATION NUMBER: US 60/131,971  
; PRIOR FILING DATE: 1999-04-30  
; PRIOR APPLICATION NUMBER: US 60/132,065  
; PRIOR FILING DATE: 1999-04-30  
; PRIOR APPLICATION NUMBER: US 60/143,928  
; PRIOR FILING DATE: 1999-07-14  
; PRIOR APPLICATION NUMBER: US 60/145,915  
; PRIOR FILING DATE: 1999-07-27  
; PRIOR APPLICATION NUMBER: US 60/146,453  
; PRIOR FILING DATE: 1999-07-29  
; PRIOR APPLICATION NUMBER: US 60/146,452  
; PRIOR FILING DATE: 1999-07-29  
; PRIOR APPLICATION NUMBER: US 60/162,288  
; PRIOR FILING DATE: 1999-10-28  
; PRIOR APPLICATION NUMBER: US 09/416,384  
; PRIOR FILING DATE: 1999-10-12  
; NUMBER OF SEQ ID NOS: 231  
; SOFTWARE: Patent.pm  
; SEQ ID NO 222  
; LENGTH: 3001  
; TYPE: DNA  
; ORGANISM: Homo Sapiens  
; FEATURE:  
; NAME/KEY: allele  
; LOCATION: 1501  
; OTHER INFORMATION: 99-27349-267 : polymorphic base G or A  
; FEATURE:  
; NAME/KEY: misc binding  
; LOCATION: 1502..1521  
; OTHER INFORMATION: 99-27349-267.misl, complement  
; FEATURE:  
; NAME/KEY: misc binding  
; LOCATION: 1482..1500

```
; OTHER INFORMATION: 99-27349-267.mis2
; FEATURE:
; NAME/KEY: primer_bind
; LOCATION: 1748..1767
; OTHER INFORMATION: upstream amplification primer, complement
; FEATURE:
; NAME/KEY: primer_bind
; LOCATION: 1337..1355
; OTHER INFORMATION: downstream amplification primer
; FEATURE:
; NAME/KEY: misc binding
; LOCATION: 1489..1513
; OTHER INFORMATION: 99-27349-267 probe
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 182,848,1501,2206,2397
; OTHER INFORMATION: n=a, g, c or t
US-09-539-333D-222

Query Match          0.6%; Score 24; DB 4; Length 3001;
Best Local Similarity 100.0%; Pred. No. 0.92;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3803 TGTCAAAAAAAAAAAAAAAAAAAAAA 3826
Db 2367 TGTCAAAAAAAAAAAAAAAAAAAAAA 2390

RESULT 2
US-09-112-096-28
; Sequence 28, Application US/09112096
; Patent No. 6194152
; GENERAL INFORMATION:
; APPLICANT: Reiner Laus
; APPLICANT: Michael H. Shapero
; APPLICANT: Larisa Tsavaler
; TITLE OF INVENTION: Prostate Tumor Polynucleotide and
; FILE REFERENCE: 7636-0015.30
; CURRENT APPLICATION NUMBER: US/09/112,096
; CURRENT FILING DATE: 1998-07-09
; EARLIER APPLICATION NUMBER: 60/056,110
; EARLIER FILING DATE: 1997-08-20
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 28
; LENGTH: 3848
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-112-096-28

Query Match          0.6%; Score 24; DB 3; Length 3848;
Best Local Similarity 100.0%; Pred. No. 0.91;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3803 TGTCAAAAAAAAAAAAAAAAAAAAAA 3826
Db 3780 TGTCAAAAAAAAAAAAAAAAAAAAAA 3803

RESULT 3
US-08-381-691-16/c
; Sequence 16, Application US/08381691
; Patent No. 5852224
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Alpha-lac Albumin Gene Constructs
; NUMBER OF SEQUENCES: 17
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
```

```
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/381,691
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3952 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-381-691-16

Query Match          0.6%; Score 24; DB 2; Length 3952;
Best Local Similarity 100.0%; Pred. No. 0.91;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3803 TGTCAAAAAAAAAAAAAAAAAAAAAA 3826
Db 3856 TGTCAAAAAAAAAAAAAAAAAAAAAA 3833

RESULT 4
US-09-144-367-1
; Sequence 1, Application US/09144367
; Patent No. 6432639
; GENERAL INFORMATION:
; APPLICANT: Lichter, Jay
; APPLICANT: Guido, Marco
; TITLE OF INVENTION: GENOTYPING OF HUMAN CYP3A4
; FILE REFERENCE: SEQ-12P
; CURRENT APPLICATION NUMBER: US/09/144,367
; CURRENT FILING DATE: 1998-08-31
; PRIOR APPLICATION NUMBER: 60/058,612
; PRIOR FILING DATE: 1997-09-10
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 2759
; TYPE: DNA
; ORGANISM: H. sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (70)...(1581)
; OTHER INFORMATION: Human CYP3A4 cDNA reference sequence
US-09-144-367-1

Query Match          0.6%; Score 23; DB 4; Length 2759;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3804 GTCAAAAAAAAAAAAAAAAAAAAAA 3826
Db 2726 GTCAAAAAAAAAAAAAAAAAAAAAA 2748

RESULT 5
US-09-705-299-10/c
; Sequence 10, Application US/09705299
; Patent No. 6440737
; GENERAL INFORMATION:
; APPLICANT: Lex M. Cowser
; APPLICANT: Susan M. Preier
; TITLE OF INVENTION: ANTISENSE MODULATION OF CELLULAR APOPTOSIS SUSCEPTIBILITY GENE
; FILE REFERENCE: RTS-0174
; CURRENT APPLICATION NUMBER: US/09/705,299
; CURRENT FILING DATE: 2000-11-01
; NUMBER OF SEQ ID NOS: 86
; SEQ ID NO 10
; LENGTH: 2929
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-705-299-10
```

Query Match 0.6%; Score 23; DB 4; Length 2929;  
 Best Local Similarity 100.0%; Pred. No. 2.5;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3802 ATGCAAAAAAAAAAAAAAAAAAAAA 3824  
 |||||  
 Db 1806 ATGCAAAAAAAAAAAAAAAAAAAAA 1784

## RESULT 6

US-08-123-934A-5  
 ; Sequence 5, Application US/08123934A  
 ; Patent No. 6291206  
 ; GENERAL INFORMATION:  
 ; APPLICANT: WOZNEY, John  
 ; APPLICANT: CELESTE, Anthony J.  
 ; APPLICANT: THIES, R. Scott  
 ; APPLICANT: YAMAJI, No. 6291206oru  
 ; TITLE OF INVENTION: RECEPTOR PROTEINS  
 ; NUMBER OF SEQUENCES: 19  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Genetics Institute Inc.- Legal Affairs  
 ; STREET: 87 CambridgePark Drive  
 ; CITY: Cambridge  
 ; STATE: MA  
 ; COUNTRY: USA  
 ; ZIP: 02140  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/123,934A  
 ; FILING DATE: 17-SEP-1993  
 ; CLASSIFICATION: 530  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: LAZAR, Steven R  
 ; REGISTRATION NUMBER: 32,618  
 ; REFERENCE/DOCKET NUMBER: 5203  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 617 876 1170  
 ; TELEFAX: 617 876 5851  
 ; INFORMATION FOR SEQ ID NO: 5:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 3238 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: DNA (genomic)  
 ; IMMEDIATE SOURCE:  
 ; CLONE: CPK1-10a  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: 474..2000  
 ; US-08-123-934A-5

Query Match 0.6%; Score 23; DB 3; Length 3238;  
 Best Local Similarity 100.0%; Pred. No. 2.5;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3804 GTCAAAAAAAAAAAAAAAAAAAAA 3826  
 |||||  
 Db 3157 GTCAAAAAAAAAAAAAAAAAAAAA 3179

## RESULT 7

US-09-874-628-5  
 ; Sequence 5, Application US/09874628  
 ; Patent No. 6610513  
 ; GENERAL INFORMATION:  
 ; APPLICANT: WOZNEY, John

CELESTE, Anthony J.  
 THIES, R. Scott  
 YAMAJI, No. 6610513oru  
 TITLE OF INVENTION: RECEPTOR PROTEINS  
 NUMBER OF SEQUENCES: 19  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Genetics Institute Inc.- Legal Affairs  
 STREET: 87 CambridgePark Drive  
 CITY: Cambridge  
 STATE: MA  
 COUNTRY: USA  
 ZIP: 02140  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/874,628  
 FILING DATE: 05-Jun-2001  
 CLASSIFICATION: <Unknown>  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/123,934  
 FILING DATE: 17-SEP-1993  
 ATTORNEY/AGENT INFORMATION:  
 NAME: LAZAR, Steven R  
 REGISTRATION NUMBER: 32,618  
 REFERENCE/DOCKET NUMBER: 5203  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 617 876 1170  
 TELEFAX: 617 876 5851  
 INFORMATION FOR SEQ ID NO: 5:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 3238 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: DNA (genomic)  
 IMMEDIATE SOURCE:  
 CLONE: CPK1-10a  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 474..2000  
 SEQUENCE DESCRIPTION: SEQ ID NO: 5:  
 US-09-874-628-5

Query Match 0.6%; Score 23; DB 4; Length 3238;  
 Best Local Similarity 100.0%; Pred. No. 2.5;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3804 GTCAAAAAAAAAAAAAAAAAAAAA 3826  
 |||||  
 Db 3157 GTCAAAAAAAAAAAAAAAAAAAAA 3179

## RESULT 8

PCT-US94-10080-5  
 ; Sequence 5, Application PC/TUS9410080  
 ; GENERAL INFORMATION:  
 ; APPLICANT: GENETICS INSTITUTE, INC.  
 ; TITLE OF INVENTION: RECEPTOR PROTEINS  
 ; NUMBER OF SEQUENCES: 19  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Genetics Institute Inc.- Legal Affairs  
 ; STREET: 87 CambridgePark Drive  
 ; CITY: Cambridge  
 ; STATE: MA  
 ; COUNTRY: USA  
 ; ZIP: 02140  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS

```

; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/10080
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/123,934
; FILING DATE: 17-SEP-1993
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: LAZAR, Steven R
; REGISTRATION NUMBER: 32,618
; REFERENCE/DOCKET NUMBER: 5203-PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 498-8260
; TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3238 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; IMMEDIATE SOURCE:
; CLONE: CFK1-10a
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 474..2000
PCT-US94-10080-5

Query Match 0.6%; Score 23; DB 5; Length 3238;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3804 GTCACAAAAA...AAAAAAAAA 3826
Db 3157 GTCACAAAAA...AAAAAAAAA 3179

RESULT 9
US-09-688-188B-1
; Sequence 1, Application US/09698188B
; Patent No. 6656716
; GENERAL INFORMATION:
; APPLICANT: FLOWMAN, GREGORY
; APPLICANT: MARTINEZ, RICARDO
; APPLICANT: WHYTE, DAVID
; TITLE OF INVENTION: STE20-RELATED PROTEIN KINASES
; CURRENT APPLICATION NUMBER: US/09/688,188B
; PRIOR FILING DATE: 2000-10-16
; PRIOR APPLICATION NUMBER: 09/291,417
; PRIOR FILING DATE: 1999-04-14
; PRIOR APPLICATION NUMBER: 60/081,784
; PRIOR FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 3268
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-688-188B-1

Query Match 0.6%; Score 23; DB 4; Length 3268;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3804 GTCACAAAAA...AAAAAAAAA 3826
Db 3214 GTCACAAAAA...AAAAAAAAA 3236

RESULT 10
US-09-819-989-1
; Sequence 1, Application US/09819989
; Sequence 1, Application US/09291417D
; Patent No. 6680170
; GENERAL INFORMATION:
; APPLICANT: FLOWMAN, GREGORY
; APPLICANT: MARTINEZ, RICARDO
; APPLICANT: WHYTE, DAVID
; TITLE OF INVENTION: STE20-RELATED PROTEIN KINASES
; CURRENT APPLICATION NUMBER: US/09/291,417D
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: 60/081,784
; PRIOR FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 3268
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-291-417D-1

Query Match 0.6%; Score 23; DB 4; Length 3268;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3804 GTCACAAAAA...AAAAAAAAA 3826
Db 3214 GTCACAAAAA...AAAAAAAAA 3236

RESULT 11
US-09-336-643A-82
; Sequence 82, Application US/09336643A
; Patent No. 6399761
; GENERAL INFORMATION:
; APPLICANT: Miller, Andrew P.
; APPLICANT: Curran, Mark Edward
; APPLICANT: Hu, Ping
; APPLICANT: Rutter, Marc
; APPLICANT: Wang, Jian-Wang
; TITLE OF INVENTION: No. 6399761el Human Potassium Channels
; FILE REFERENCE: SEQ-15P
; CURRENT APPLICATION NUMBER: US/09/336,643A
; CURRENT FILING DATE: 1999-06-18
; PRIOR APPLICATION NUMBER: 60/076,687
; PRIOR FILING DATE: 1998-08-07
; PRIOR APPLICATION NUMBER: 60/116,448
; PRIOR FILING DATE: 1999-01-19
; PRIOR APPLICATION NUMBER: PCT/US99/03826
; PRIOR FILING DATE: 1999-02-22
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 82
; LENGTH: 3300
; TYPE: DNA
; ORGANISM: H. sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (50)...(1285)
US-09-336-643A-82

Query Match 0.6%; Score 23; DB 4; Length 3300;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3804 GTCACAAAAA...AAAAAAAAA 3826
Db 3242 GTCACAAAAA...AAAAAAAAA 3264

RESULT 12
US-09-819-989-1
; Sequence 1, Application US/09819989
```

```
; Patent No. 6482629
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al.
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
; FILE REFERENCE: THEREOF
; FILE REFERENCE: CL001200
; CURRENT APPLICATION NUMBER: US/09/819,989
; CURRENT FILING DATE: 2001-03-29
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3377
; TYPE: DNA
; ORGANISM: Human
US-09-819-989-1

Query Match      0.6%; Score 23; DB 4; Length 3377;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3804 GTCACAAAAA 3826
Db 3339 GTCACAAAAA 3361

RESULT 13
US-10-273-992-1
; Sequence 1, Application US/10273992
; Patent No. 6664093
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al.
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
; FILE REFERENCE: THEREOF
; FILE REFERENCE: CL001200DIV
; CURRENT APPLICATION NUMBER: US/10/273,992
; CURRENT FILING DATE: 2002-10-21
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3377
; TYPE: DNA
; ORGANISM: Human
US-10-273-992-1

Query Match      0.6%; Score 23; DB 4; Length 3377;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3804 GTCACAAAAA 3826
Db 3339 GTCACAAAAA 3361

RESULT 14
US-09-866-028-6
; Sequence 6, Application US/09866028
; Patent No. 6642360
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Botstein, David
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Geritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul
; APPLICANT: Grimaldi, Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Hillan, Kenneth
; APPLICANT: Kijavini, Ivar
; APPLICANT: Napier, Mary
```

```
; APPLICANT: Roy, Margaret
; APPLICANT: Tumas, Daniel
; APPLICANT: Wood, William
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P2548P1C1
; CURRENT APPLICATION NUMBER: US/09/866,028
; CURRENT FILING DATE: 2001-05-25
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 120
; SEQ ID NO 6
; LENGTH: 3441
; TYPE: DNA
; ORGANISM: Homo Sapien
US-09-866-028-6

Query Match      0.6%; Score 23; DB 4; Length 3441;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3804 GTCACAAAAA 3826
Db 3406 GTCACAAAAA 3428

RESULT 15
US-09-083-351-1
; Sequence 1, Application US/09083351
; Patent No. 6087107
; GENERAL INFORMATION:
; APPLICANT: Sheffield, Val C.
; APPLICANT: Alward, Wallace L.M.
; APPLICANT: Stone, Edwin M.
; APPLICANT: Nishimura, Darryl
; APPLICANT: Patil, Shiva
; TITLE OF INVENTION: THERAPEUTICS AND DIAGNOSTICS FOR
; TITLE OF INVENTION: CONGENITAL HEART DISEASE BASED ON A NOVEL HUMAN
; TITLE OF INVENTION: TRANSCRIPTION FACTOR
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/083,351
; FILING DATE: 22-MAY-1998
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Arnold, Beth E.
; REGISTRATION NUMBER: 35,430
; REFERENCE/DOCKET NUMBER: UIA-029.02
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-832-1000
; TELEFAX: 617-832-7000
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3946 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 475..2133
US-09-083-351-1
```

Query Match 0.6%; Score 23; DB 3; Length 3946;  
 Best Local Similarity 100.0%; Pred. No. 2.4;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3804 GTCAAAAAAAAAAAAAAAAAAAA 3826  
 |||||  
 Db 3919 GTCAAAAAAAAAAAAAAAAAAAA 3941

Search completed: September 4, 2004, 13:14:42  
 Job time : 185 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 4, 2004, 08:32:53 ; Search time 1152 Seconds  
(without alignments)  
16508.391 Million cell updates/sec

Title: US-09-927-091-3

Perfect score: 3826

Sequence: 1 aggtcgctggaccgaagc.....aaaaaaaaaaaaaaaaaa 3826

Scoring table: OLIGO\_NUC

Gapop 60.0 , Gapext 60.0

Searched: 3267054 seqs, 2485319735 residues

Word size : 18

Total number of hits satisfying chosen parameters: 32167

Minimum DB seq length: 2500

Maximum DB seq length: 5000

Post-processing: Listing first 45 summaries

Database : Published Applications NA:\*

1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*

2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*

3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*

4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*

5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*

6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*

7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*

8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*

9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq:\*

10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq:\*

11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*

12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*

13: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*

14: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq:\*

15: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq:\*

16: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq:\*

17: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*

18: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*

19: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3826	100.0	3826	9	US-09-927-091-3
2	26	0.7	4702	16	US-10-295-027-141
3	26	0.7	4702	16	US-10-295-027-1164
4	25	0.7	2511	17	US-10-437-963-100661
5	25	0.7	2522	9	US-09-745-763-39
6	25	0.7	2586	16	US-10-104-047-875
7	25	0.7	2629	15	US-10-106-698-1916
8	25	0.7	2796	9	US-09-529-063-51
9	25	0.7	2796	15	US-10-414-378-51
10	25	0.7	2806	13	US-10-424-599-137631
11	25	0.7	3465	17	US-10-615-659-11
12	25	0.7	3504	17	US-10-437-963-54840
13	25	0.7	3554	17	US-10-615-659-12
14	25	0.7	4572	17	US-10-450-826-108

15	25	0.7	4740	9	US-09-945-265-3	Sequence 3, Appli
16	25	0.7	4740	10	US-09-902-481A-2	Sequence 2, Appli
17	25	0.7	4740	15	US-10-207-655-175	Sequence 175, App
18	25	0.7	4740	15	US-10-136-819-18	Sequence 18, Appl
19	25	0.7	4740	16	US-10-159-563-339	Sequence 339, App
20	24	0.6	2509	9	US-09-925-301-540	Sequence 540, App
21	24	0.6	2602	17	US-10-437-963-11080	Sequence 11080, A
22	24	0.6	2633	9	US-09-764-877-2154	Sequence 2154, Ap
23	24	0.6	2633	16	US-10-242-515-2154	Sequence 799, App
24	24	0.6	2645	14	US-10-044-090-799	Sequence 399, App
25	24	0.6	2661	16	US-10-159-563-399	Sequence 4, Appli
26	24	0.6	2661	17	US-10-211-179-4	Sequence 1, Appli
27	24	0.6	2677	10	US-09-733-643-1	Sequence 611, App
28	24	0.6	2840	9	US-09-978-295A-611	Sequence 611, App
29	24	0.6	2840	9	US-09-978-697-611	Sequence 611, App
30	24	0.6	2840	9	US-09-978-192A-611	Sequence 611, App
31	24	0.6	2840	9	US-09-999-832A-611	Sequence 611, App
32	24	0.6	2840	10	US-09-978-189-611	Sequence 611, App
33	24	0.6	2840	10	US-09-978-608A-611	Sequence 611, App
34	24	0.6	2840	10	US-09-978-585A-611	Sequence 611, App
35	24	0.6	2840	10	US-09-978-191A-611	Sequence 611, App
36	24	0.6	2840	10	US-09-978-403A-611	Sequence 611, App
37	24	0.6	2840	10	US-09-978-564A-611	Sequence 611, App
38	24	0.6	2840	10	US-09-999-833A-611	Sequence 611, App
39	24	0.6	2840	10	US-09-981-915A-611	Sequence 611, App
40	24	0.6	2840	10	US-09-978-824-611	Sequence 611, App
41	24	0.6	2840	10	US-09-918-585A-611	Sequence 611, App
42	24	0.6	2840	10	US-09-978-423A-611	Sequence 611, App
43	24	0.6	2840	10	US-09-978-193A-611	Sequence 611, App
44	24	0.6	2840	10	US-09-999-830A-611	Sequence 611, App
45	24	0.6	2840	10	US-09-978-757A-611	Sequence 611, App

## ALIGNMENTS

RESULT 1

US-09-927-091-3

; Sequence 3, Application US/09927091

; Patent No. US20020119541A1

; GENERAL INFORMATION:

; APPLICANT: KILLARY, ANN

; APPLICANT: LOTT, STEVE

; APPLICANT: CHANDLER, DAWN

; TITLE OF INVENTION: THE TUMOR SUPPRESSOR CAR-1

; FILE REFERENCE: US/09/927,091

; CURRENT APPLICATION NUMBER: US/09/927,091

; PRIOR FILING DATE: 2001-08-09

; PRIOR APPLICATION NUMBER: 60/227,560

; PRIOR FILING DATE: 2000-08-23

; PRIOR APPLICATION NUMBER: 60/225,033

; PRIOR FILING DATE: 2000-08-10

; NUMBER OF SEQ ID NOS: 9

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 3

; LENGTH: 3826

; TYPE: DNA

; ORGANISM: Human

US-09-927-091-3

Query Match	100.0%;	Score 3826;	DB 9;	Length 3826;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 3826;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	AGCTCGCTGGACCGAAGCGGTGGCTGAAGCTCCGCGGGGTAAAGGGTCCGCTGGG	60	
Db	1	AGCTCGCTGGACCGAAGCGGTGGCTGAAGCTCCGCGGGGTAAAGGGTCCGCTGGG	60	
QY	61	CCAGGGTTTGGGCGCGGATCCGGAGCTGAGGGGGCGGCACCCCTCTCTCTGTC	120	
Db	61	CCAGGGTTTGGGCGCGGATCCGGAGCTGAGGGGGCGGCACCCCTCTCTCTGTC	120	
QY	121	GGTCACAGCAATGTACGGCTCGGCTGGCTGCCCTCCCTCCAGGATTCCCATCCCCA	180	

121	Db		GGTCACAGCCAAATGTACGGCTCGGCTGGCTGGCCCTCCCCAGAGATTCCCATCCCCA	180
181	Qy		GCTTCTCGCCTCCCGCACGGCCCCACCCCGGATTTTCGACCCCTTAAGGGCTCCAC	240
181	Db		GCTTCTCGCCTCCCGCACGGCCCCACCCCGGATTTTCGACCCCTTAAGGGCTCCAC	240
241	Qy		CCGCTCCGGATCCCTTCTCCAGCTCCTATCCTCCCTTAGGATGTCGCCGCCCCCTAGAA	300
241	Db		CCGCTCCGGATCCCTTCTCCAGCTCCTATCCTCCCTTAGGATGTCGCCGCCCCCTAGAA	300
301	Qy		CCTCCCGTCAAGATCTCCGTCCTCAGCGCTCAGAGCTCTCTCCAGCGCCATCGCC	360
301	Db		CCTCCCGTCAAGATCTCCGTCCTCAGCGCTCAGAGCTCTCTCCAGCGCCATCGCC	360
361	Qy		TTGAGCTGCCCACTACCTCTAGATGCCCTCCCGGCTGGCTGCCAGGAGTCTCAGCC	420
361	Db		TTGAGCTGCCCACTACCTCTAGATGCCCTCCCGGCTGGCTGCCAGGAGTCTCAGCC	420
421	Qy		GCACACCCCTTCTCGGCTTACCTCTTCGGACAGCACCCCTCTCCCTTCCGGTAGC	480
421	Db		GCACACCCCTTCTCGGCTTACCTCTTCGGACAGCACCCCTCTCCCTTCCGGTAGC	480
481	Qy		TCTTACCCCTCTGTGTGGGCTGTGTCCCGCGCCGAGCCCTGTGCTGTGCCTCGACA	540
481	Db		TCTTACCCCTCTGTGTGGGCTGTGTCCCGCGCCGAGCCCTGTGCTGTGCCTCGACA	540
541	Qy		GGCGCGGCTCTCTCAGCGCCCCCTCGCCCTGGCGCCGCCCTCTCTGCTGCCCTGGC	600
541	Db		GGCGCGGCTCTCTCAGCGCCCCCTCGCCCTGGCGCCGCCCTCTCTGCTGCCCTGGC	600
601	Qy		GCCATGGCTGAGCCTCAAGACGAGTGTGTGTCTCATCTGCCTGAGCATCTACCA	660
601	Db		GCCATGGCTGAGCCTCAAGACGAGTGTGTGTCTCATCTGCCTGAGCATCTACCA	660
661	Qy		GACCCGCTGAGCTGGCTCGAGCACTACTTCTCGCGCGCTGCATCAGGAGCACTGG	720
661	Db		GACCCGCTGAGCTGGCTCGAGCACTACTTCTCGCGCGCTGCATCAGGAGCACTGG	720
721	Qy		GTGCGGAGGAGGCGAGGGGCGCCCGCATGCCCCAGTGGCGGCGCATGTCGCGAG	780
721	Db		GTGCGGAGGAGGCGAGGGGCGCCCGCATGCCCCAGTGGCGGCGCATGTCGCGAG	780
781	Qy		CCGCGCTGGCGCCAGCCTCAAGCTGGCAACATGTGGAGCGCTACAGTCTCTTCCCG	840
781	Db		CCGCGCTGGCGCCAGCCTCAAGCTGGCAACATGTGGAGCGCTACAGTCTCTTCCCG	840
841	Qy		CTGGAAGCATCTCAAGCGCGCGCGCGCGACCTTGCCAGGCGCACGAAGGTC	900
841	Db		CTGGAAGCATCTCAAGCGCGCGCGCGCGACCTTGCCAGGCGCACGAAGGTC	900
901	Qy		AAGCTTCTGCTCAGCGACCGCGCTTCTGCTTCTTCTGCGAGGCGCTGACCTG	960
901	Db		AAGCTTCTGCTCAGCGACCGCGCTTCTGCTTCTTCTGCGAGGCGCTGACCTG	960
961	Qy		CACGAGCAGCATCAGGTCAACCGCATCGACAGCCCTTCAGCAGCTGCGAGGAGCTG	1020
961	Db		CACGAGCAGCATCAGGTCAACCGCATCGACAGCCCTTCAGCAGCTGCGAGGAGCTG	1020
1021	Qy		AAGGACCAACTTCAGGCGCTTCAAGACAGCGAGCGGGAACACACCGAAGCGCTGAG	1080
1021	Db		AAGGACCAACTTCAGGCGCTTCAAGACAGCGAGCGGGAACACACCGAAGCGCTGAG	1080
1081	Qy		CTAAGCGACAACCTGGCGGAGACCAAGTCTTCCACCAAGAGCCTGCGGACCACTATCGG	1140
1081	Db		CTAAGCGACAACCTGGCGGAGACCAAGTCTTCCACCAAGAGCCTGCGGACCACTATCGG	1140
1141	Qy		GAGGCTTCGAGCGGCTGCACCGGCTGTGCGTGAACGCCAGAGGCGCATGCTAGAGAG	1200
1141	Db		GAGGCTTCGAGCGGCTGCACCGGCTGTGCGTGAACGCCAGAGGCGCATGCTAGAGAG	1200
1201	Qy		CTGGAAGGCGGACACGGGCCGACCGTGCACCGACATCGAGCAGAAAGTCCAGCGCTACAGC	1260

Db	1201	CTGAGCGCGGACACGGCCCGCAGCTGTACCGGACATGAGGACAGAAAGTCCAGCGCTACAGC	1261
Qy	1261	CAGCAGCTGGCGCAAGGTCCAGGAGGGAGCCAGATCCTGCAGAGCGGTGGCTGGAAC	1320
Db	1261	CAGCAGCTGGCGCAAGGTCCAGGAGGGAGCCAGATCCTGCAGAGCGGTGGCTGGAAC	1320
Qy	1321	GACCGGCACACCTTCTGGCTGGGTGGCTCTACTGTCCGAGCGGTCAAGGGAAATC	1380
Db	1321	GACCGGCACACCTTCTGGCTGGGTGGCTCTACTGTCCGAGCGGTCAAGGGAAATC	1380
Qy	1381	CATGAGACCAACTCTACATATGAAGACTTCCCGACCTCCAAAGTACACAGGCCCTCTGCAG	1440
Db	1381	CATGAGACCAACTCTACATATGAAGACTTCCCGACCTCCAAAGTACACAGGCCCTCTGCAG	1440
Qy	1441	TACACCATCTGGAAAGTCCCTGTTCCAGGACATCCACCAGTGGCAGCGCCCTTAACCCCTG	1500
Db	1441	TACACCATCTGGAAAGTCCCTGTTCCAGGACATCCACCAGTGGCAGCGCCCTTAACCCCTG	1500
Qy	1501	GACCCGGGCACAGCCACAGCGCTGATCCTGTGCGACGACTGCACCATGTGGGTTCAC	1560
Db	1501	GACCCGGGCACAGCCACAGCGCTGATCCTGTGCGACGACTGCACCATGTGGGTTCAC	1560
Qy	1561	GGCAACTTGACCCACAGCCACTGCAGGACTTCGCGCAAGCGCTTCGATGTGGAGGTGTGCG	1620
Db	1561	GGCAACTTGACCCACAGCCACTGCAGGACTTCGCGCAAGCGCTTCGATGTGGAGGTGTGCG	1620
Qy	1621	GTGCTGGGTCTGGAAGCCTTCAGTAGTGGCGTCCACTACTGCGAGGTGGTGGCGGGAG	1680
Db	1621	GTGCTGGGTCTGGAAGCCTTCAGTAGTGGCGTCCACTACTGCGAGGTGGTGGCGGGAG	1680
Qy	1681	AAGACCCAGTGGGTGATCGGGCTGGCACAGAAAGCGCGCAAGCGCGCAAGCGCGCATCCAG	1740
Db	1681	AAGACCCAGTGGGTGATCGGGCTGGCACAGAAAGCGCGCAAGCGCGCGCATCCAG	1740
Qy	1741	ATCCAGCCGACGCGCGCTTCTACTGATCGTGATGCAAGTGGCAACGATGGCAACGATCAGCGCC	1800
Db	1741	ATCCAGCCGACGCGCGCTTCTACTGATCGTGATGCAAGTGGCAACGATGGCAACGATCAGCGCC	1800
Qy	1801	TGCACGGAGCCCTGCGACGGCGCTTAAGCTCCGGGACAAGCTTGCAAGGTGGGTGTCTTC	1860
Db	1801	TGCACGGAGCCCTGCGACGGCGCTTAAGCTCCGGGACAAGCTTGCAAGGTGGGTGTCTTC	1860
Qy	1861	CTGGACTATGACCAAGGCTTGTCTCATCTTCTACAAATGCTGATGACATGTCTGTGCTCTAC	1920
Db	1861	CTGGACTATGACCAAGGCTTGTCTCATCTTCTACAAATGCTGATGACATGTCTGTGCTCTAC	1920
Qy	1921	ACCTTCGCGAGAGTTCCTGCGCAAGCTGTCTTACTTTCAGCCCTGCGCAGAGCCAC	1980
Db	1921	ACCTTCGCGAGAGTTCCTGCGCAAGCTGTCTTACTTTCAGCCCTGCGCAGAGCCAC	1980
Qy	1981	GCCAAATGGCAGAAAGTTCAGCGCGCTCGGATCAACACCGTCCGATCTAGTCCAGCGAG	2040
Db	1981	GCCAAATGGCAGAAAGTTCAGCGCGCTCGGATCAACACCGTCCGATCTAGTCCAGCGAG	2040
Qy	2041	AAGAGACCAACCTCTCTGGGACCACTGCCACTGCAAGAGCCCTGCCAGGAAGATG	2100
Db	2041	AAGAGACCAACCTCTCTGGGACCACTGCCACTGCAAGAGCCCTGCCAGGAAGATG	2100
Qy	2101	AAGACCTGGACTCCAGGCCACCGTGGCCACTGGAGAGCTCAGGCCAGTGTTCACCCCTCC	2160
Db	2101	AAGACCTGGACTCCAGGCCACCGTGGCCACTGGAGAGCTCAGGCCAGTGTTCACCCCTCC	2160
Qy	2161	AGCCTCCAGTGTGAAATGAGGTTCGATTCCTCTTCTGTAACCTCTTTCAGCATC	2220
Db	2161	AGCCTCCAGTGTGAAATGAGGTTCGATTCCTCTTCTGTAACCTCTTTCAGCATC	2220
Qy	2221	GATGTTCTGTAGTCTGACCTTGTATAGGATACAGCTTTGATCCAAAGATGTGACATGGC	2280
Db	2221	GATGTTCTGTAGTCTGACCTTGTATAGGATACAGCTTTGATCCAAAGATGTGACATGGC	2280
Qy	2281	TTCTCCTCAGGGCAACCCCTGCCCAACCCCTCATCCCCATCTTCTCAGGGCAGGGGACTA	2340
Db	2281	TTCTCCTCAGGGCAACCCCTGCCCAACCCCTCATCCCCATCTTCTCAGGGCAGGGGACTA	2340

QY 2341 CTTTCCAGTGTCTCCCTCCAGCCAGCCCTGACCTCAGGAAGTGTAGAGCATGGCCAGT 2400  
DB 2341 CTTTCCAGTGTCTCCCTCCAGCCAGCCCTGACCTCAGGAAGTGTAGAGCATGGCCAGT 2400  
QY 2401 AGTTGGCAGCCCGAAAGACACACAGCACCTCTTTATGTCCCATGGCCCTAAGACTTACCCC 2460  
DB 2401 AGTTGGCAGCCCGAAAGACACACAGCACCTCTTTATGTCCCATGGCCCTAAGACTTACCCC 2460  
QY 2461 TGACCAAGCTAGTGTATGGGCAATTTACCTTGACCCAGTCCACAGTGGTCCACAGTAGT 2520  
DB 2461 TGACCAAGCTAGTGTATGGGCAATTTACCTTGACCCAGTCCACAGTGGTCCACAGTAGT 2520  
QY 2521 ACCTGGTCCCTAGGTTGCTCCAGAGCAACCTCTCTGCGCCACCCACACCAAGAACTAT 2580  
DB 2521 ACCTGGTCCCTAGGTTGCTCCAGAGCAACCTCTCTGCGCCACCCACACCAAGAACTAT 2580  
QY 2581 ATGGTTCCCTACTTCTCCACTGATCTGCTGGTCAAGTGTGATGCTGTGGCTGTGGAAGG 2640  
DB 2581 ATGGTTCCCTACTTCTCCACTGATCTGCTGGTCAAGTGTGATGCTGTGGCTGTGGAAGG 2640  
QY 2641 CACTGCTAGTGTGAGTCCACACATTAAGTATGATGTCACACCTTCTCTGCGCCACAGGCC 2700  
DB 2641 CACTGCTAGTGTGAGTCCACACATTAAGTATGATGTCACACCTTCTCTGCGCCACAGGCC 2700  
QY 2701 GAGGACAGGGTGTAGGGTATACCCAAAGCTGTGTCAGAGCCCATTAAGCTTAAAGCAACT 2760  
DB 2701 GAGGACAGGGTGTAGGGTATACCCAAAGCTGTGTCAGAGCCCATTAAGCTTAAAGCAACT 2760  
QY 2761 GCAGGACAGCCCTCCCTGGATGATCGAGGTCCCGAGTGTCTGAAACAGAGTCCAGCCA 2820  
DB 2761 GCAGGACAGCCCTCCCTGGATGATCGAGGTCCCGAGTGTCTGAAACAGAGTCCAGCCA 2820  
QY 2821 ACCCTCTTACAGCAGGCTCTGACCTGTAGCTGTAGGTCAGAGGCTTCCAGAGCAGTTG 2880  
DB 2821 ACCCTCTTACAGCAGGCTCTGACCTGTAGCTGTAGGTCAGAGGCTTCCAGAGCAGTTG 2880  
QY 2881 TTGTAATTAGACCCCAAGACTGGAGGGGCTGTGGTGTAGACCCCTGTGACAGTTGGC 2940  
DB 2881 TTGTAATTAGACCCCAAGACTGGAGGGGCTGTGGTGTAGACCCCTGTGACAGTTGGC 2940  
QY 2941 ATCTATCTCAGTTAGTATCTGCTGCAAGAAACAAAGAGCCACTGTAGCTGGTTTAATTA 3000  
DB 2941 ATCTATCTCAGTTAGTATCTGCTGCAAGAAACAAAGAGCCACTGTAGCTGGTTTAATTA 3000  
QY 3001 GACAAGGATTTACTACTGCGCCCTGTGCTGTGCAAAATTTGTAAGAGCTGGAGAAG 3060  
DB 3001 GACAAGGATTTACTACTGCGCCCTGTGCTGTGCAAAATTTGTAAGAGCTGGAGAAG 3060  
QY 3061 CAGACTCTGTAATTTCCAGGAACCTCCAGCGCCAGATTCATCTCTGTTGTGACCA 3120  
DB 3061 CAGACTCTGTAATTTCCAGGAACCTCCAGCGCCAGATTCATCTCTGTTGTGACCA 3120  
QY 3121 GGAAGAGTCCGCCATCTGCAAGAACCACTATGCCAGAAAGCTGTGACTGCAGAACTA 3180  
DB 3121 GGAAGAGTCCGCCATCTGCAAGAACCACTATGCCAGAAAGCTGTGACTGCAGAACTA 3180  
QY 3181 GGCCTCCTCTGCCACGGTCCGTCAGCCCAATAGATGCTCTGAGGCTGCCCTCTCCCA 3240  
DB 3181 GGCCTCCTCTGCCACGGTCCGTCAGCCCAATAGATGCTCTGAGGCTGCCCTCTCCCA 3240  
QY 3241 CTTCACTCAGTCTCCCAATCTAAATTTTACAGAGATTTCTGTTGGGGAACCTTAAGTC 3300  
DB 3241 CTTCACTCAGTCTCCCAATCTAAATTTTACAGAGATTTCTGTTGGGGAACCTTAAGTC 3300  
QY 3301 AGATCCAGAACCTTGGTGTCAAGGGAGTCTGGGAAATGTCAATTTCCCTAGAAGGAAGTTA 3360  
DB 3301 AGATCCAGAACCTTGGTGTCAAGGGAGTCTGGGAAATGTCAATTTCCCTAGAAGGAAGTTA 3360  
QY 3361 GGGTGGGTGAGCAAGCCCAACCTGGTGTCTGCGCACAGCATCCATCGTGAAGAACT 3420  
DB 3361 GGGTGGGTGAGCAAGCCCAACCTGGTGTCTGCGCACAGCATCCATCGTGAAGAACT 3420

QY 3421 CGGAGAGGCTGGAGTCCACATCTAGGGTTGTCTCTGCCCTTGCTCTATCCCTGCCAG 3480  
DB 3421 CGGAGAGGCTGGAGTCCACATCTAGGGTTGTCTCTGCCCTTGCTCTATCCCTGCCAG 3480  
QY 3481 AGTGGGAACCTGGAGGAGTGGGCTGCAAGACTGAGCCCTAAATGTCTCCCGGCTTGACT 3540  
DB 3481 AGTGGGAACCTGGAGGAGTGGGCTGCAAGACTGAGCCCTAAATGTCTCCCGGCTTGACT 3540  
QY 3541 TTTCTTTCTAGTCTCTGGGCTAGATTTCTGCACCTTGGGGTCTCTGACACACACCATC 3600  
DB 3541 TTTCTTTCTAGTCTCTGGGCTAGATTTCTGCACCTTGGGGTCTCTGACACACACCATC 3600  
QY 3601 CCAAGTAGCCGGAAGACTAAACACAGGGGGTCTTAAATGGCTGCCCGCCACCCG 3660  
DB 3601 CCAAGTAGCCGGAAGACTAAACACAGGGGGTCTTAAATGGCTGCCCGCCACCCG 3660  
QY 3661 GGCTCCCTTGGGCAAGAAATTTTATTTAAATGTCAGCCCTACCCCAACCTTCAACCAATCT 3720  
DB 3661 GGCTCCCTTGGGCAAGAAATTTTATTTAAATGTCAGCCCTACCCCAACCTTCAACCAATCT 3720  
QY 3721 GGCCCAACCCAGCAGTATTTTATTTAAATGTCAGCCCTTGGCCATTTATGATGATCAAT 3780  
DB 3721 GGCCCAACCCAGCAGTATTTTATTTAAATGTCAGCCCTTGGCCATTTATGATGATCAAT 3780  
QY 3781 TGTATTAAATTAAGTTACAGATGTCAAAAAAAGGCTCAAAAAAAGGCTCAAAAAAAGG 3826  
DB 3781 TGTATTAAATTAAGTTACAGATGTCAAAAAAAGGCTCAAAAAAAGGCTCAAAAAAAGG 3826

## RESULT 2

US-10-295-027-141  
; Sequence 141, Application US/10295027  
; Publication No. US20030232350A1  
; GENERAL INFORMATION:  
; APPLICANT: Afar, Daniel  
; APPLICANT: Aziz, Natasha  
; APPLICANT: Ginsberg, Wendy M.  
; APPLICANT: Gish, Kurt C.  
; APPLICANT: Glynn, Richard  
; APPLICANT: Hevezi, Peter A.  
; APPLICANT: Mack, David H.  
; APPLICANT: Murray, Richard  
; APPLICANT: Watson, Susan R.  
; APPLICANT: Eos Biotechnology, Inc.  
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and  
; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer  
; FILE REFERENCE: 018501-012500US  
; CURRENT APPLICATION NUMBER: US/10/295,027  
; CURRENT FILING DATE: 2002-11-13  
; PRIOR APPLICATION NUMBER: US 09/663,733  
; PRIOR FILING DATE: 2000-09-15  
; PRIOR APPLICATION NUMBER: US 60/350,666  
; PRIOR FILING DATE: 2001-11-13  
; PRIOR APPLICATION NUMBER: US 60/335,394  
; PRIOR FILING DATE: 2001-11-15  
; PRIOR APPLICATION NUMBER: US 60/332,464  
; PRIOR FILING DATE: 2001-11-21  
; PRIOR APPLICATION NUMBER: US 60/334,393  
; PRIOR FILING DATE: 2001-11-29  
; PRIOR APPLICATION NUMBER: US 60/340,376  
; PRIOR FILING DATE: 2001-12-14  
; PRIOR APPLICATION NUMBER: US 60/347,211  
; PRIOR FILING DATE: 2002-01-08  
; PRIOR APPLICATION NUMBER: US 60/347,349  
; PRIOR FILING DATE: 2002-01-10  
; PRIOR APPLICATION NUMBER: US 60/355,250  
; PRIOR FILING DATE: 2002-02-08  
; PRIOR APPLICATION NUMBER: US 60/356,714  
; PRIOR FILING DATE: 2002-02-13  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 1386  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 141

```
; LENGTH: 4702
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-295-027-141

Query Match          0.7%; Score 26; DB 16; Length 4702;
Best Local Similarity 100.0%; Pred. No. 0.049;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3801 GATGTCACAAAAA 3826
      |||||
Db 4660 GATGTCACAAAAA 4685

RESULT 3
US-10-295-027-1164
; Sequence 1164, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1164
; LENGTH: 4702
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-295-027-1164

Query Match          0.7%; Score 26; DB 16; Length 4702;
Best Local Similarity 100.0%; Pred. No. 0.049;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3801 GATGTCACAAAAA 3826
      |||||
Db 4660 GATGTCACAAAAA 4685

RESULT 4
US-10-437-963-100661/c
; Sequence 100661, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 100661
; LENGTH: 2511
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_98356C.1
US-10-437-963-100661

Query Match          0.7%; Score 25; DB 17; Length 2511;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3802 ATGTCACAAAAA 3826
      |||||
Db 26 ATGTCACAAAAA 2

RESULT 5
US-09-745-763-39
; Sequence 39, Application US/09745763
; Patent No. US20020065394A1
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; McCoy, John M.
; Lavallie, Edward R.
; Collins-Racie, Lisa A.
; Evans, Cheryl
; Merberg, David
; Treacy, Maurice
; Spaulding, Vikki
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
; ENCODING THEM
; NUMBER OF SEQUENCES: 219
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 Cambridgepark Drive
; CITY: Cambridge
; STATE: MA
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/745,763
; FILING DATE: 18-Jun-2000
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Sprunger, Suzanne A.
; REGISTRATION NUMBER: 41,323
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 498-8284
; TELEFAX: (617) 876-5851
```

```
/ INFORMATION FOR SEQ ID NO: 39:
/ SEQUENCE CHARACTERISTICS:
/   LENGTH: 2522 base pairs
/   TYPE: nucleic acid
/   STRANDEDNESS: double
/   TOPOLOGY: linear
/   MOLECULE TYPE: cDNA
/   SEQUENCE DESCRIPTION: SEQ ID NO: 39:
US-09-745-763-39

Query Match          0.7%; Score 25; DB 9; Length 2522;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3802 ATGTCACAAAAA 3826
Db 2495 ATGTCACAAAAA 2519

RESULT 6
US-10-104-047-875/c
/ Sequence 875, Application US/10104047
/ Publication No. US20030236392A1
/ GENERAL INFORMATION:
/ APPLICANT: HELIX RESEARCH INSTITUTE
/ TITLE OF INVENTION: No. US20030236392A1el full length cDNA
/ FILE REFERENCE: HI-A0105
/ CURRENT APPLICATION NUMBER: US/10/104,047
/ CURRENT FILING DATE: 2002-03-25
/ PRIOR APPLICATION NUMBER:
/ PRIOR FILING DATE:
/ NUMBER OF SEQ ID NOS: 4096
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 875
/ LENGTH: 2586
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/
US-10-104-047-875

Query Match          0.7%; Score 25; DB 16; Length 2586;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3802 ATGTCACAAAAA 3826
Db 2112 ATGTCACAAAAA 2088

RESULT 7
US-10-106-698-1916
/ Sequence 1916, Application US/10106698
/ Publication No. US20030109690A1
/ GENERAL INFORMATION:
/ APPLICANT: Ruben et al.
/ TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
/ FILE REFERENCE: PA005P1
/ CURRENT APPLICATION NUMBER: US/10/106,698
/ CURRENT FILING DATE: 2002-03-27
/ PRIOR APPLICATION NUMBER: PCT/US00/26524
/ PRIOR FILING DATE: 2000-09-28
/ PRIOR APPLICATION NUMBER: US 60/157,137
/ PRIOR FILING DATE: 1999-09-29
/ PRIOR APPLICATION NUMBER: US 60/163,280
/ PRIOR FILING DATE: 1999-11-03
/ NUMBER OF SEQ ID NOS: 8564
/ SOFTWARE: PatentIn Ver. 3.0
/ SEQ ID NO 1916
/ LENGTH: 2629
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (35)..(35)
```

```
/ OTHER INFORMATION: n equals a,t,g, or c
US-10-106-698-1916

Query Match          0.7%; Score 25; DB 15; Length 2629;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3802 ATGTCACAAAAA 3826
Db 2567 ATGTCACAAAAA 2591

RESULT 8
US-09-529-063-51
/ Sequence 51, Application US/09529063
/ Patent No. US20020102542A1
/ GENERAL INFORMATION:
/ APPLICANT: FUKUSHIMA, DAIKICHI
/ APPLICANT: SHIBAYAMA, SHIRO
/ APPLICANT: TADA, HIDEAKI
/ TITLE OF INVENTION: POLYPEPTIDE, cDNA ENCODING THE POLYPEPTIDE, AND USE OF
/ TITLE OF INVENTION: THE BOTH
/ FILE REFERENCE: Q58769
/ CURRENT APPLICATION NUMBER: US/09/529,063
/ CURRENT FILING DATE: 2000-04-07
/ PRIOR APPLICATION NUMBER: PCT/JP98/04514
/ PRIOR FILING DATE: 1998-10-06
/ PRIOR APPLICATION NUMBER: JP 9-274674
/ PRIOR FILING DATE: 1997-10-07
/ NUMBER OF SEQ ID NOS: 117
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 51
/ LENGTH: 2796
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: sig_peptide
/ LOCATION: (11)..(58)
/ NAME/KEY: mat_peptide
/ LOCATION: (59)..(1273)
/ NAME/KEY: CDS
/ LOCATION: (11)..(1273)
/
US-09-529-063-51

Query Match          0.7%; Score 25; DB 9; Length 2796;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3802 ATGTCACAAAAA 3826
Db 2768 ATGTCACAAAAA 2792

RESULT 9
US-10-414-378-51
/ Sequence 51, Application US/10414378
/ Publication No. US20030165981A1
/ GENERAL INFORMATION:
/ APPLICANT: FUKUSHIMA, DAIKICHI
/ APPLICANT: SHIBAYAMA, SHIRO
/ APPLICANT: TADA, HIDEAKI
/ TITLE OF INVENTION: POLYPEPTIDE, cDNA ENCODING THE POLYPEPTIDE, AND USE OF
/ TITLE OF INVENTION: THE BOTH
/ FILE REFERENCE: Q58769
/ CURRENT APPLICATION NUMBER: US/10/414,378
/ CURRENT FILING DATE: 2003-04-16
/ PRIOR APPLICATION NUMBER: US/09/529,063
/ PRIOR FILING DATE: 2000-04-07
/ PRIOR APPLICATION NUMBER: PCT/JP98/04514
/ PRIOR FILING DATE: 1998-10-06
/ PRIOR APPLICATION NUMBER: JP 9-274674
/ PRIOR FILING DATE: 1997-10-07
/ NUMBER OF SEQ ID NOS: 117
```

; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 51  
; LENGTH: 2796  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: sig\_peptide  
; LOCATION: (11)..(58)  
; FEATURE:  
; NAME/KEY: mat\_peptide  
; LOCATION: (59)..(1273)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (11)..(1273)  
; US-10-414-378-51

Query Match 0.7%; Score 25; DB 15; Length 2796;  
Best Local Similarity 100.0%; Pred. No. 0.16;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3802 ATGTCAAAAAAAAAAAAAAAAAAAAA 3826  
|||||  
DB 2768 ATGTCAAAAAAAAAAAAAAAAAAAAA 2792

RESULT 10  
US-10-424-599-137631  
; Sequence 137631, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 137631  
; LENGTH: 2806  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_9528C.1  
US-10-424-599-137631

Query Match 0.7%; Score 25; DB 13; Length 2806;  
Best Local Similarity 100.0%; Pred. No. 0.16;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3802 ATGTCAAAAAAAAAAAAAAAAAAAAA 3826  
|||||  
DB 2455 ATGTCAAAAAAAAAAAAAAAAAAAAA 2479

RESULT 11  
US-10-615-659-11  
; Sequence 11, Application US/10615659  
; Publication No. US2004015723A1  
; GENERAL INFORMATION:  
; APPLICANT: Bristol-Myers Squibb Company  
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING A NOVEL TESTIS-SPECIFIC TUBULIN  
; TITLE OF INVENTION: TYROSINE-LIGASE-LIKE PROTEIN, BGS42  
; FILE REFERENCE: D0283 NP  
; CURRENT APPLICATION NUMBER: US/10/615,659  
; CURRENT FILING DATE: 2003-07-09  
; PRIOR APPLICATION NUMBER: U.S. 60/394,725  
; PRIOR FILING DATE: 2002-07-09  
; NUMBER OF SEQ ID NOS: 102  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 11

; LENGTH: 3465  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; US-10-615-659-11

Query Match 0.7%; Score 25; DB 17; Length 3465;  
Best Local Similarity 100.0%; Pred. No. 0.16;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3802 ATGTCAAAAAAAAAAAAAAAAAAAAA 3826  
|||||  
DB 3441 ATGTCAAAAAAAAAAAAAAAAAAAAA 3465

RESULT 12  
US-10-437-963-54840  
; Sequence 54840, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 54840  
; LENGTH: 3504  
; TYPE: DNA  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_56905C.1  
US-10-437-963-54840

Query Match 0.7%; Score 25; DB 17; Length 3504;  
Best Local Similarity 100.0%; Pred. No. 0.16;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3802 ATGTCAAAAAAAAAAAAAAAAAAAAA 3826  
|||||  
DB 3467 ATGTCAAAAAAAAAAAAAAAAAAAAA 3491

RESULT 13  
US-10-615-659-12  
; Sequence 12, Application US/10615659  
; Publication No. US2004015723A1  
; GENERAL INFORMATION:  
; APPLICANT: Bristol-Myers Squibb Company  
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING A NOVEL TESTIS-SPECIFIC TUBULIN  
; TITLE OF INVENTION: TYROSINE-LIGASE-LIKE PROTEIN, BGS42  
; FILE REFERENCE: D0283 NP  
; CURRENT APPLICATION NUMBER: US/10/615,659  
; CURRENT FILING DATE: 2003-07-09  
; PRIOR APPLICATION NUMBER: U.S. 60/394,725  
; PRIOR FILING DATE: 2002-07-09  
; NUMBER OF SEQ ID NOS: 102  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 12  
; LENGTH: 3554  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; US-10-615-659-12

Query Match 0.7%; Score 25; DB 17; Length 3554;  
Best Local Similarity 100.0%; Pred. No. 0.16;

```
Query Match          0.7%; Score 25; DB 9; Length 4740;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```



**This Page Blank (uspto)**

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 4, 2004, 08:24:02 ; Search time 6095 Seconds  
(without alignments)  
18745.316 Million cell updates/sec

Title: US-09-927-091-3  
Perfect score: 3826  
Sequence: 1 aggcgcctggaccgaagc.....aaaaaaaaaaaaaaaaaaaaa 3826

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 27513289 seqs, 14931090276 residues

Word size : 18  
Total number of hits satisfying chosen parameters: 3147

Minimum DB seq length: 2500  
Maximum DB seq length: 5000

Post-processing: Listing first 45 summaries

Database :

- EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_htc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_htc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_man:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rod:\*  
26: em\_gss\_pbg:\*  
27: em\_gss\_vri:\*  
28: gb\_gss1:\*  
29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	119	3.1	2726	11 AK048725	AK048725 Mus muscu
2	25	0.7	3080	11 BC037619	BC037619 Mus muscu
3	24	0.6	2779	11 BC052223	BC052223 Danio rer
4	24	0.6	2963	11 AK085089	AK085089 Mus muscu

5	24	0.6	3234	11 AK081017	AK081017 Mus muscu
6	24	0.6	3250	11 BC035530	BC035530 Mus muscu
7	24	0.6	3415	14 CA916724	CA916724 K58FL hum
8	24	0.6	3444	11 BC039604	BC039604 Homo sapi
9	24	0.6	3707	13 C83838	C83838 C3838 Dict
10	24	0.6	4097	11 BC050053	BC050053 Homo sapi
11	24	0.6	4168	11 BC042403	BC042403 Mus muscu
12	24	0.6	4207	11 BC034315	BC034315 Homo sapi
13	23	0.6	2669	11 BC035179	BC035179 Homo sapi
14	23	0.6	2752	11 BC014615	BC014615 Homo sapi
15	23	0.6	2802	11 BC030311	BC030311 Mus muscu
16	23	0.6	2819	11 BC051011	BC051011 Mus muscu
17	23	0.6	2860	11 BC043612	BC043612 Homo sapi
18	23	0.6	2907	11 BC024884	BC024884 Mus muscu
19	23	0.6	2911	11 AK090089	AK090089 Mus muscu
20	23	0.6	2942	11 AK090144	AK090144 Mus muscu
21	23	0.6	2970	11 BC023335	BC023335 Mus muscu
22	23	0.6	2978	11 BC030337	BC030337 Mus muscu
23	23	0.6	3066	11 AK040164	AK040164 Mus muscu
24	23	0.6	3089	11 AK079444	AK079444 Mus muscu
25	23	0.6	3091	11 BC036748	BC036748 Homo sapi
26	23	0.6	3109	11 AK039198	AK039198 Mus muscu
27	23	0.6	3266	11 BC035538	BC035538 Mus muscu
28	23	0.6	3301	11 AK029761	AK029761 Mus muscu
29	23	0.6	3395	11 BC033458	BC033458 Mus muscu
30	23	0.6	3444	11 BC008551	BC008551 Mus muscu
31	23	0.6	3466	11 AK052111	AK052111 Mus muscu
32	23	0.6	3741	11 BC034973	BC034973 Homo sapi
33	23	0.6	3814	11 BC016415	BC016415 Mus muscu
34	23	0.6	4042	11 BC052704	BC052704 Mus muscu
35	23	0.6	4585	11 BC059828	BC059828 Mus muscu
36	23	0.6	4617	11 BC007668	BC007668 Homo sapi
37	23	0.6	4619	11 BC058338	BC058338 Mus muscu
38	22	0.6	2504	11 BC011059	BC011059 Mus muscu
39	22	0.6	2507	11 AK048762	AK048762 Mus muscu
40	22	0.6	2509	11 BC036220	BC036220 Homo sapi
41	22	0.6	2512	11 BC036441	BC036441 Homo sapi
42	22	0.6	2520	11 AK086585	AK086585 Mus muscu
43	22	0.6	2521	11 BC049779	BC049779 Mus muscu
44	22	0.6	2527	11 BC042136	BC042136 Homo sapi
45	22	0.6	2528	11 AF451985	AF451985 Homo sapi

ALIGNMENTS

RESULT 1  
AK048725  
LOCUS  
DEFINITION

AK048725 2726 bp mRNA linear HTC 20-SEP-2003  
Mus musculus 0 day neonate cerebellum cDNA, RIKEN full-length  
enriched library, clone:C230023D16 product:similar to CDNA FLJ10759  
FIS, CLONE NT2RP3004617, WEAKLY SIMILAR TO ZINC-BINDING PROTEIN A33  
[Homo sapiens], full insert sequence.

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

AK048725.1 GI:26093070  
HTC; CAP trapper.  
Mus musculus (house mouse)  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
PUBMED

1  
Carninci, P. and Hayashizaki, Y.  
High-efficiency full-length cDNA cloning  
Meth. Enzymol. 303, 19-44 (1999)  
99279253  
10349636

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE

2  
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new genes  
Genome Res. 10 (10), 1617-1630 (2000)  
20499374

```

PUBMED
REFERENCE
AUTHORS
11042159
3 Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
Konno,H., Akiyama,J., Nishi,K., Kitsuai,T., Tashiro,H., Itoh,M.,
Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
Yanamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
Fujiwaki,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
Okazaki,Y., Muramatsu,M., Inoue,Y., Kita,A. and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multipillar sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
JOURNAL
MEDLINE
11076861
4
The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 2726)
Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
Okazaki,Y., Saioh,K., Saioh,H., Sakai,C., Sakai,K., Sakazume,N.,
Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
Segabe,Y., Tagami,M., Tagawa,A., Tanaka,T., Tanaka,A., Takahashi,F., Takaku-Akahira,S.,
Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
Muramatsu,M. and Hayashizaki,Y.
Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/.
FEATURES
source
Location/Qualifiers
1. .2726
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:C230023D16"
/db_xref="MG1:2415428"
/db_xref="taxon:10090"
/clone="C230023D16"
/tissue_type="cerebellum"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="0 day neonate"
1. .2726
/note="similar to CDNA FLJ10759 FIS, CLONE NT2RP3004617,
WEAKLY SIMILAR TO ZINC-BINDING PROTEIN A33 [Homo sapiens]
(SFTR|Q9NVG0, evidence: FASTY, 74.8%ID, 61.4%length,
match=1154)"
misc_feature
1. .2726
/note="similar to CDNA FLJ10759 FIS, CLONE NT2RP3004617,
WEAKLY SIMILAR TO ZINC-BINDING PROTEIN A33 [Homo sapiens]
(SFTR|Q9NVG0, evidence: FASTY, 74.8%ID, 61.4%length,
match=1154)"
ORIGIN
Query Match 0.7%; Score 25; DB 11; Length 3080;
Best Local Similarity 100.0%; Pred. No. 1.4e+04;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps C;

Query Match 3.1%; Score 119; DB 11; Length 2726;
Best Local Similarity 100.0%; Pred. No. 3.6e+14;
Matches 119; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 592 GCCCTGGCGCATTGGCTGAGCTCAAGACGAGCTGCTGCTCATCTGCTGAGC 651
|||||
Db 590 GCCCTGGCGCATTGGCTGAGCTCAAGACGAGCTGCTGCTCATCTGCTGAGC 639
|||||
QY 652 ATCTACAGGACCCGGTGGCTGGCTGGGAGCAGCTACTTCTCCGCCGCTGCATCAC 710
|||||
Db 640 ATCTACAGGACCCGGTGGCTGGCTGGGAGCAGCTACTTCTCCGCCGCTGCATCAC 698
|||||

RESULT 2
BC037619
LOCUS BC037619 3080 bp mRNA linear HTC 04-MAR-2003
DEFINITION Mus musculus, clone IMAGE:5400144, mRNA.
ACCESSION BC037619
VERSION BC037619.1 GI:22902286
KEYWORDS HTC.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Strausberg,R.
Direct Submission
Submitted (13-SEP-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: gcapsb-remail.nih.gov
Tissue Procurement: The Cepko Laboratory
cDNA Library Prepared by: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: http://www.nisc.nih.gov/
Contact: nisc_mgc@nhgri.nih.gov
Akhtar,N., Ayale,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R.,
Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C.,
McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W.,
Tsurgueon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,
Young,A., Zhang,L.-H. and Green,E.D.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 81 Row: p Column: 13
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis, GenomeScan gene prediction
This clone has the following problem: retained intron.
FEATURES
source
Location/Qualifiers
1. .3080
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:5400144"
/tissue_type="Eye, retina, mouse strain C57BL/6"
/clone_lib="NIH_MGC_94"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
ORIGIN
Query Match 0.7%; Score 25; DB 11; Length 3080;
Best Local Similarity 100.0%; Pred. No. 1.4e+04;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps C;

```

REFERENCE	AUTHORS
TITLE	JOURNAL
REFERENCE	AUTHORS
TITLE	

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAK Plate: 104 Row: n Column: 23  
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis. Similarity but not identity to protein

JOURNAL REFERENCE  
AUTHORS

Nature 420, 563-573 (2002)  
6 (bases 1 to 2963)  
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

TITLE  
JOURNAL

Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL: http://genome.gsc.riken.go.jp/ URL: http://fantom.gsc.riken.go.jp/

FEATURES  
source

Location/Qualifiers  
1..2963  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="FANTOM:DB:D430035E22"  
/db\_xref="MGI:2421960"  
/db\_xref="taxon:10090"  
/clone="D430035E22"  
/tissue\_type="lung"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="13 days embryo"  
1..2963  
/notes="unclassifiable"

misc\_feature

ORIGIN

Query Match 0.6%; Score 24; DB 11; Length 2963;  
Best Local Similarity 100.0%; Pred. No. 2.3e+04;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3803 TGTCAAAAAAAAAAAAAAAAAAAAAA 3826  
|||||  
Db 24 TGTCAAAAAAAAAAAAAAAAAAAAAA 1

RESULT 5  
AK081017 LOCUS  
DEFINITION  
Mus musculus 10 days neonate cerebellum cDNA, RIKEN full-length enriched library, clone:B930054008 product:unclassifiable, full insert sequence.

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

AK081017  
AK081017.1 GI:26099613  
HTC; CAP trapper.  
Mus musculus (house mouse)  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE

Carninci, P. and Hayashizaki, Y.  
High-efficiency full-length cDNA cloning  
Meth. Enzymol. 303, 19-44 (1999)  
99279253

10349636  
2  
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes  
Genome Res. 10 (10), 1617-1630 (2000)  
20499374  
11042159  
3  
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuina, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M., Itonaka, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer  
Genome Res. 10 (11), 1757-1771 (2000)  
20530913  
11076861  
4  
The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.  
Functional annotation of a full-length mouse cDNA collection  
Nature 409, 685-690 (2001)  
5  
The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)  
6 (bases 1 to 3234)  
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.  
Direct Submission  
Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)  
Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL: http://genome.gsc.riken.go.jp/ URL: http://fantom.gsc.riken.go.jp/

FEATURES  
source

Location/Qualifiers  
1..3234  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="FANTOM:DB:B930054008"  
/db\_xref="MGI:2412933"  
/db\_xref="taxon:10090"  
/clone="B930054008"  
/tissue\_type="cerebellum"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"



## ORIGIN

Query Match 0.6%; Score 24; DB 14; Length 3415;  
 Best Local Similarity 100.0%; Pred. No. 2e+04;  
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3803 TGTCAAAAAAAAAAAAAAAAAAAAAA 3826

Db 3388 TGTCAAAAAAAAAAAAAAAAAAAAAA 3411

## RESULT 8

BC039604

LOCUS BC039604 3444 bp mRNA linear HTC 04-MAR-2003

DEFINITION Homo sapiens, clone IMAGE:3847010, mRNA.

ACCESSION BC039604

VERSION BC039604.1 GI:25058528

KEYWORDS HTC.

SOURCE Homo sapiens (human)

ORGANISM

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

TITLE 1 (bases 1 to 3444)

JOURNAL Strausberg, R.

Submitted (01-NOV-2002) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>

COMMENT Contact: MGC help desk

Email: [cgabs@mail.nih.gov](mailto:cgabs@mail.nih.gov)

Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA sequencing by: Institute for Systems Biology

<http://www.systemsbio.org>

contact: [amad@systemsbiology.org](mailto:amad@systemsbiology.org)

Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha

Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Series: IRAK Plate: 84 Row: a Column: 13  
 This clone was selected for full length sequencing because it passed the following selection criteria: GenomeScan gene prediction  
 This clone has the following problem: retained intron.

## FEATURES

source

1..3444  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:3847010"  
 /tissue\_type="Colon, adenocarcinoma"  
 /clone\_lib="NIH MGC\_65"  
 /lab\_host="DH10B"  
 /note="Vector: pCMV-SPORT6"

## ORIGIN

Query Match 0.6%; Score 24; DB 11; Length 3444;  
 Best Local Similarity 100.0%; Pred. No. 2e+04;  
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3803 TGTCAAAAAAAAAAAAAAAAAAAAAA 3826

Db 3419 TGTCAAAAAAAAAAAAAAAAAAAAAA 3442

## RESULT 9

CB8388

LOCUS

DEFINITION CB8388 Dictyostelium discoideum SS (H.Urushihara) Dictyostelium

discoideum cDNA clone SSA355, mRNA sequence.

## ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

COMMENT

10048482

Contact: Hideko Urushihara

Institute of Biological Sciences

University of Tsukuba

1-1-1 Tennoudai, Tsukuba, Ibaraki 305-8572, Japan

Tel: 81-298-53-4664

Fax: 81-298-53-6614

Email: [hideko@biol.tsukuba.ac.jp](mailto:hideko@biol.tsukuba.ac.jp)

PROJECT = 'Dictyostelium discoideum cDNA project in Japan'

POLYA=No.

Location/Qualifiers

1..3707

/organism="Dictyostelium discoideum"

/mol\_type="mRNA"

/strain="AX4"

/db\_xref="taxon:44689"

/clone="SSA355"

/dev\_stage="slug"

/clone\_lib="Dictyostelium discoideum SS (H.Urushihara)"

ORIGIN

Query Match 0.6%; Score 24; DB 13; Length 3707;

Best Local Similarity 100.0%; Pred. No. 1.9e+04;

Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3803 TGTCAAAAAAAAAAAAAAAAAAAAAA 3826

Db 3632 TGTCAAAAAAAAAAAAAAAAAAAAAA 3655

RESULT 10

BC050053

LOCUS

DEFINITION

Homo sapiens DKFP434I092 protein, mRNA (cDNA clone IMAGE:483990),

containing frame-shift errors.

ACCESSION BC050053

VERSION BC050053.1 GI:29747709

KEYWORDS HTC.

SOURCE Homo sapiens (human)

ORGANISM

REFERENCE

AUTHORS

1 (bases 1 to 4097)

Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,

Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,

Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,

Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,

Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,

Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,

Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,

Carninci, P., Prange, C., Raja, S.S., Loquellano, N.A., Peters, G.J.,

Abrahamson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J.,

McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,

Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,

Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,

Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodrigues, S.,

Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,

Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,



Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smailus, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.  
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences  
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
22388257  
12477932  
2 (bases 1 to 4097)  
Strausberg, R.  
Direct Submission  
Submitted (01-APR-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA  
NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgaps-remail.nih.gov](mailto:cgaps-remail.nih.gov)  
cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki Toshiyuki and Piero Carninci (RIKEN)  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center  
Center code: BCM-HGSC  
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>  
Contact: [amg@bcm.tmc.edu](mailto:amg@bcm.tmc.edu)  
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H., Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati, A.N., Gibbs, R.A.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAX Plate: 72 Row: P Column: 7  
This clone has the following problem: frame shifted.  
This clone has the following problem: frame shifted.  
Location/Qualifiers  
1. .4097  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4839390"  
/tissue\_type="Testis"  
/clone\_lib="NIH MGC\_97"  
/lab\_host="DH10B"  
/note="Vector: pBluescript"

FEATURES  
source  
Query Match 0.6%; Score 24; DB 11; Length 4097;  
Best Local Similarity 100.0%; Pred. No. 1.7e+04;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3803 TGTCACAAAAA 3826  
Db 4029 TGTCAAAAAA 4052

ORIGIN  
RESULT 11  
BC042403 4168 bp mRNA linear HTC 19-NOV-2003  
LOCUS  
DEFINITION  
Mus musculus RIKEN cDNA 1200014D22 gene, mRNA (cdna clone IMAGE:3156629).  
ACCESSION  
BC042403  
VERSION  
BC042403.1 GI:27469394  
KEYWORDS  
HTC.  
SOURCE  
Mus musculus (house mouse)  
ORGANISM  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
REFERENCE  
1 (bases 1 to 4168)  
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,

Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Schetz, T.E., Brange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smailus, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.  
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences  
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
22388257  
12477932  
2 (bases 1 to 4168)  
Strausberg, R.  
Direct Submission  
Submitted (02-JAN-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA  
NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgaps-remail.nih.gov](mailto:cgaps-remail.nih.gov)  
Tissue Procurement: Jeffrey Green M.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center  
Center code: BCM-HGSC  
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>  
Contact: [amg@bcm.tmc.edu](mailto:amg@bcm.tmc.edu)  
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H., Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati, A.N., Gibbs, R.A.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAX Plate: 61 Row: b Column: 10  
This clone has the following problem: no 5' EST match.  
This clone has the following problem: no 5' EST match.  
Location/Qualifiers  
1. .4168  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="FVB/N"  
/db\_xref="taxon:10090"  
/clone="IMAGE:3156629"  
/tissue\_type="Mammary tumor. C3(1)-Tag model. Infiltrating ductal carcinoma. 5 month old virgin mouse."  
/clone\_lib="NCI CGAP\_Mam6"  
/lab\_host="DH10B"  
/note="Vector: pCMV-SPORT6"

FEATURES  
source  
Query Match 0.6%; Score 24; DB 11; Length 4168;  
Best Local Similarity 100.0%; Pred. No. 1.7e+04;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3803 TGTCACAAAAA 3826  
Db 4091 TGTCAAAAAA 4114

ORIGIN  
RESULT 12  
BC034315  
LOCUS  
DEFINITION  
Homo sapiens hypothetical protein LOC90529, mRNA (cdna clone IMAGE:4827425), containing frame-shift errors.

Qy	3803	TGTCACAAAAA	3826
Db	2376	TGTCAAAAAAAAAA	2399
RESULT 13			
BC035179			
LOCUS	Homo sapiens, clone IMAGE:5266192,	mRNA	linear HTC 04-MAR-2003
DEFINITION	BC035179		
ACCESSION	BC035179.1	GI:23273366	
VERSION			
KEYWORDS	HTC.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	Strausberg R.		
JOURNAL	Submitted (31-JUL-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA		
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov		
COMMENT	Contact: MGC help desk Email: cgaps-remail.nih.gov Tissue Procurement: Miklos Palkovits, M.D., Ph.D. CDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki Toshituki and Piero Carninci (RIKEN) CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Institute for Systems Biology http://www.systemsbiology.org contact: amadan@systemsbiology.org Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting		
Clone distribution:	MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov		
Series:	IRAK Plate: 73 Row: o Column: 3		
This clone has the following problem:	retained intron.		
FEATURES			
source	1..2669		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:5266192"		
	/tissue_type="Testis"		
	/clone_lib="NIH_MGC_97"		
	/lab_host="DH10B"		
	/note="vector: pbluescript"		
ORIGIN			
Query Match	0.6%;	Score 23;	DB 11; Length 2669;
Best Local Similarity	100.0%;	Pred. No. 3.8e+04;	
Matches	23;	Conservative 0;	Mismatches 0; Indels 0; Gaps 0;
Qy	3804	GTCACAAAAA	3826
Db	1546	GTCACAAAAA	1568
RESULT 14			
BC014615			
LOCUS	Homo sapiens cDNA clone IMAGE:3532298,	mRNA	linear HTC 17-DEC-2003
DEFINITION	BC014615		
ACCESSION	BC014615.2	GI:33869935	
VERSION			
KEYWORDS	HTC.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
JOURNAL	Submitted (31-JUL-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA		
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov		
COMMENT	Contact: MGC help desk Email: cgaps-remail.nih.gov Tissue Procurement: Miklos Palkovits, M.D., Ph.D. CDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki Toshituki and Piero Carninci (RIKEN) CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Institute for Systems Biology http://www.systemsbiology.org contact: amadan@systemsbiology.org Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting		
Clone distribution:	MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov		
Series:	IRAK Plate: 73 Row: o Column: 3		
This clone has the following problem:	retained intron.		
FEATURES			
source	1..2669		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:5266192"		
	/tissue_type="Testis"		
	/clone_lib="NIH_MGC_97"		
	/lab_host="DH10B"		
	/note="vector: pbluescript"		
ORIGIN			
Query Match	0.6%;	Score 23;	DB 11; Length 2669;
Best Local Similarity	100.0%;	Pred. No. 3.8e+04;	
Matches	23;	Conservative 0;	Mismatches 0; Indels 0; Gaps 0;
Qy	3804	GTCACAAAAA	3826
Db	1546	GTCACAAAAA	1568
RESULT 14			
BC014615			
LOCUS	Homo sapiens cDNA clone IMAGE:3532298,	mRNA	linear HTC 17-DEC-2003
DEFINITION	BC014615		
ACCESSION	BC014615.2	GI:33869935	
VERSION			
KEYWORDS	HTC.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
JOURNAL	Submitted (31-JUL-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA		
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov		
COMMENT	Contact: MGC help desk Email: cgaps-remail.nih.gov Tissue Procurement: Miklos Palkovits, M.D., Ph.D. CDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki Toshituki and Piero Carninci (RIKEN) CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Institute for Systems Biology http://www.systemsbiology.org contact: amadan@systemsbiology.org Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting		
Clone distribution:	MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov		
Series:	IRAK Plate: 73 Row: o Column: 3		
This clone has the following problem:	retained intron.		
FEATURES			
source	1..2669		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:5266192"		
	/tissue_type="Testis"		
	/clone_lib="NIH_MGC_97"		
	/lab_host="DH10B"		
	/note="vector: pbluescript"		
ORIGIN			
Query Match	0.6%;	Score 23;	DB 11; Length 2669;
Best Local Similarity	100.		



```

/lab_host="DH10B"
/note="Vector: PCMV-SPORT6"

ORIGIN
Query Match      0.6%; Score 23; DB 11; Length 2802;
Best Local Similarity 100.0%; Pred. No. 3.7e+04;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      3804  GTCAAAAAAAAAAAAAAAAAAAAA 3826
          |||||||
Db      2778  GTCAAAAAAAAAAAAAAAAAAAAA 2800

Search completed: September 4, 2004, 13:11:20
Job time : 6098 secs

```